

THE



MANUAL

A practical guide to
delirium assessment

ALASDAIR MACLULLICH

THE 4AT MANUAL

A PRACTICAL GUIDE TO DELIRIUM ASSESSMENT

ALASDAIR MACLULLICH

DELIRIUM SUPPORT PRESS

First edition, 2026

Manual version 1.0

Published in Edinburgh by Delirium Support Press.

Paperback edition.

ISBN 978-1-918164-00-8.

Copyright © 2026 Alasdair MacLullich. Except where otherwise stated, the text, tables and original diagrams in this manual are licensed under the Creative Commons Attribution 4.0 International licence (CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/>. The 4AT oval logo, cover artwork and design, and separately credited third-party material are excluded. The licence does not grant trademark rights or permit any implication of endorsement.

Notwithstanding those exclusions, the publisher permits the cover artwork, cover design and 4AT oval logo to be reproduced unaltered only as integral parts of complete, unmodified copies of this edition. This permission does not permit those elements to be extracted, adapted or used separately, does not grant trademark rights, and must not be used to imply endorsement, sponsorship or official status.

The 4AT was originally designed by Alasdair MacLullich (University of Edinburgh) with Tracy Ryan and Helen Cash (NHS Lothian, Scotland). Responsibility for ongoing 4AT work is with Alasdair MacLullich. The definitive 4AT form and supporting materials are at www.the4AT.com.

Recommended citation: MacLullich AMJ. The 4AT Manual: a practical guide to delirium assessment. First edition, version 1.0. Edinburgh: Delirium Support Press; 2026. Available from: www.the4AT.com

Funding: No external funding was received for the preparation or publication of this manual.

Competing interests: Alasdair MacLullich is the principal author of the 4AT and is responsible for its ongoing development and official website. The 4AT is free to use, and he receives no licence fees from its clinical use. He may receive royalties from sales of paid editions of this manual. No other competing interests are declared.

Enquiries, corrections and permissions: a.maclullich@gmail.com · www.the4AT.com

Medical information only. This manual supports clinical assessment; it does not replace clinical judgement or individualised advice.

For my parents, Claudine and Archie, with love and gratitude.

CONTENTS

<i>Introduction</i>	v
1. Why the 4AT exists	1
2. Delirium: the essentials	7
3. Administering the 4AT: item by item	11
4. Scoring and interpretation	18
5. What to do after a positive 4AT	23
6. The 4AT across settings	27
7. Detection as a system: episodic tests and monitoring	33
8. Clinical cases	36
9. Implementing the 4AT, and knowing whether it is working	58
10. The evidence base	62
11. The 4AT compared with other tools	67
12. Versions, translations and terms of use	73
13. Related tools: the SQiD and the 4-DSD	75
14. Frequently asked questions	77
15. Resources	83
<i>References</i>	87

INTRODUCTION

ABOUT THIS MANUAL

This manual is for people who use the 4AT in clinical work: nurses, doctors, occupational therapists, physiotherapists, paramedics, pharmacists, physician associates and students. It is also for the people who make detection happen at scale: educators, quality improvement teams, audit leads and the teams who build assessment tools into electronic records. Those using the 4AT in research studies will find a detailed description of how to use the 4AT as well as comprehensive coverage of the supporting literature.

You do not need to read it in order. Most readers will want one of four things.

If you are about to do a 4AT and want the rules, go to The 4AT at a glance and Chapters 3 and 4. If you have a score and are not sure what to do with it, go to Chapters 4 and 5. If you are teaching the 4AT or checking your own practice, the cases in Chapter 8 are the fastest way in. If you are implementing the 4AT in a hospital, service or electronic record, start with Chapters 7 and 9, then use Chapter 10 when someone asks you for the evidence.

First, the 4AT indicates possible delirium, but diagnosis is a clinical act. A score of 4 or above tells the assessor that delirium may be present and that clinical assessment is now required. The same holds for every delirium instrument in existence. No tool diagnoses delirium without this clinical judgement step following tool completion.

Second, the 4AT requires no formal training course. Validation studies have generally used assessors given brief orientation to the tool rather than a training course. Even so, using any delirium tool well requires some knowledge of delirium, and familiarity with the 4AT improves consistency. Chapter 2 provides the minimum, in about ten minutes of reading. Later chapters provide more detail, including 30 case examples.

THE 4AT AT A GLANCE

The 4AT has four items, each beginning with A: Alertness, AMT₄, Attention, and Acute change or fluctuating course. It takes about two minutes. The definitive printable form and translations are at www.the4AT.com.

Item 1: Alertness. Observe the patient throughout. If asleep, attempt to wake with speech or a gentle touch on the shoulder. Ask the patient to state their name and address to assist rating. This item includes patients who are markedly drowsy (difficult to rouse, or obviously sleepy during assessment) or agitated/hyperactive.

Score 0 — Normal (fully alert, but not agitated, throughout assessment)

Score 0 — Mild sleepiness for less than 10 seconds after waking, then normal

Score 4 — Clearly abnormal

Item 2: AMT₄ (a four-item version of the Abbreviated Mental Test). Ask the patient: age, date of birth, place (name of the hospital or building), current year.

Score 0 — No mistakes

Score 1 — 1 mistake

Score 2 — 2 or more mistakes, or untestable

Item 3: Attention (Months Backwards). Ask the patient: “Please tell me the months of the year in backwards order, starting at December.” One prompt of “what is the month before December?” is permitted to assist initial understanding.

Score 0 — Achieves 7 months or more correctly

Score 1 — Starts but scores fewer than 7 months, or refuses to start

Score 2 — Untestable (cannot start because unwell, drowsy, inattentive)

Item 4: Acute change or fluctuating course. Evidence of significant change or fluctuation in alertness, cognition, or other mental function (for example paranoia or hallucinations) arising over the last two weeks and still evident in the last 24 hours. Sources: your own knowledge of the patient, other staff, letters from the general practitioner, case notes, carers and family.

Score 0 — No

Score 4 — Yes

Total score (0 to 12):

4 or above — Possible delirium, with or without cognitive impairment

1 to 3 — Possible cognitive impairment

0 — Delirium or severe cognitive impairment unlikely (delirium still possible if the Item 4 information is incomplete)

A positive score is not a diagnosis, but it is a strong reason to assess the patient clinically. And in the other direction: no patient is “unable to assess” with the 4AT. If the patient cannot engage with the

cognitive items, those items are scored untestable and the test still produces a score. This matters, because the patients who cannot engage are often the ones with delirium.

WHY THE 4AT EXISTS

Creation of the 4AT began with an audit that failed at the first patient.

In 2010 the Edinburgh team set out to measure the gap between delirium recorded in the case notes and delirium actually present on the wards, using one of the established assessment tools recommended at the time. A nurse was trained to administer it exactly as its manual required. She went to see the first patient on the audit list; MacLulich waited in the office. She was back inside two minutes. “He’s just sitting there drooling.”

The patient barely looked at her. He could not engage with conversation, let alone answer questions, and because the tool’s procedure began with cognitive testing, the assessment could not begin at all. Use of this tool had collapsed on first contact with the very kind of patient it existed to find. He was not an awkward exception. He was exactly the sort of patient the audit was meant to find, and the fact that he was drowsy without being comatose put him at the more severe end of the delirium spectrum. Yet the official process had no way to score him.

Experiences like this kept accumulating, and they indicated three

distinct types of failure of tools commonly used in routine practice. Tools often were not completed at all, and detection rates in ordinary care were dismal. When tools were completed, they could score negative in patients who plainly had delirium, usually because the required preceding cognitive testing had been skipped. And a large group of patients were being recorded as “Unable to Assess”: too drowsy or too inattentive to answer questions, and therefore, absurdly, unscorable.

That last group is the heart of the matter, and it connected directly to research the group was doing at the same time on level of arousal. In an exploratory study of acute hip fracture patients, abnormal arousal on simple bedside observation was a strong indicator of delirium [1], and this line of work later produced a dedicated observational arousal scale, the Observational Scale of Level of Arousal (OSLA) [2]. The clinical conclusion was clear: acute reduced arousal in an unwell patient nearly always signifies delirium. The same argument was made at consensus level in a position paper led by Daniel Davis with the European and American Delirium Societies, whose title says it all: “inclusiveness is safer” [3]. Patients too drowsy to speak were being excluded twice over, from diagnostic test accuracy studies and from assessment in practice, when they are not outside the delirium spectrum but at its severe end. Marking them “Unable to Assess” records the most important finding in the case as simply missing. This is serious for such patients because they would enter a kind of diagnostic limbo, with often no clear actions taken.

So the team set out to design a new tool on one principle: it had to work for the patient in front of you, rather than requiring the patient to be well enough to fit the tool. The starting question was practical. What does an experienced practitioner actually do at the bedside when quickly assessing someone for possible delirium? And can that process be turned into a short, simple instrument that anyone can use without training?

The answer had four parts, and they became the four items.

A practitioner approaching the bed first registers the patient’s level of alertness. Is she asleep, rousable, drifting off mid-sentence? Is

he agitated, hypervigilant, picking at the air? Altered alertness in a general hospital setting is highly specific to delirium, so on the 4AT a clear abnormality scores 4 on its own: enough to flag possible delirium with no further testing. Rather than being an obstacle to assessment, reduced alertness became a scoreable finding. The drooling patient from the audit scores 4 on the 4AT in the first 10–20 seconds. This is the design decision behind the rule that no patient is Unable to Assess.

Next, brief cognitive testing, built into the tool rather than requiring a separate test before the tool. Some existing instruments told the assessor to do a cognitive test without specifying which, and the completion statistics later showed what that vagueness cost. If staff are not sure what they are supposed to do, they tend to do nothing. So the 4AT names its tests and contains them.

The AMT4 (age, date of birth, place, year) is a four-item orientation test derived from the Abbreviated Mental Test, with good evidence supporting its use as a brief screen [4,5]. Months of the Year Backwards is a widely used bedside attention test, sensitive to the inattention at the core of delirium, usable with no equipment in almost any language [6,7], and consistent with the group's parallel research interest in measuring attention in delirium objectively [8].

In the 4AT, these items are scored if the patient is able to comprehend and answer the questions. Additionally, both items carry a possible untestable rating. An untestable rating for each item gives a score of 2. The untestable rating achieves two things. It turns inability to engage, itself a signal of delirium, into points on the score: untestable status on both items yields 4, flagging possible delirium. And it gives the assessor standing at the bedside of a silent patient a defined way forward, instead of stalling completion of the tool, which is what happened when the audit nurse met with the very first case.

Last, the feature that distinguishes delirium from dementia and from lifelong cognitive impairment: acute change or fluctuation. Item 4 asks whether there is evidence of significant change or fluctuation in mental function arising over the last two weeks and still present in the last 24 hours. It scores 4 when present, because this finding is

highly specific to delirium even when everything else looks equivocal.

The scoring's multiple routes to a positive result are deliberate; they reflect how the diagnosis actually gets made. Delirium may show itself through several small signs, such as a cognitive error and wandering attention, or through one overwhelming signal: this person is definitely different from their baseline. That alone is enough to act on, and the 4AT allows observation, testing, or history to carry the score across the threshold. It also means the tool still works when no informant exists, because Items 1 to 3 provide routes to 4 points on their own. Where a tool's algorithm cannot return a result until change from baseline has been established, and no informant is available, the assessment waits. Delay in a medical emergency is precisely the problem the 4AT was built to remove. The 4AT weighs acute change heavily but does not make it a precondition of producing a score.

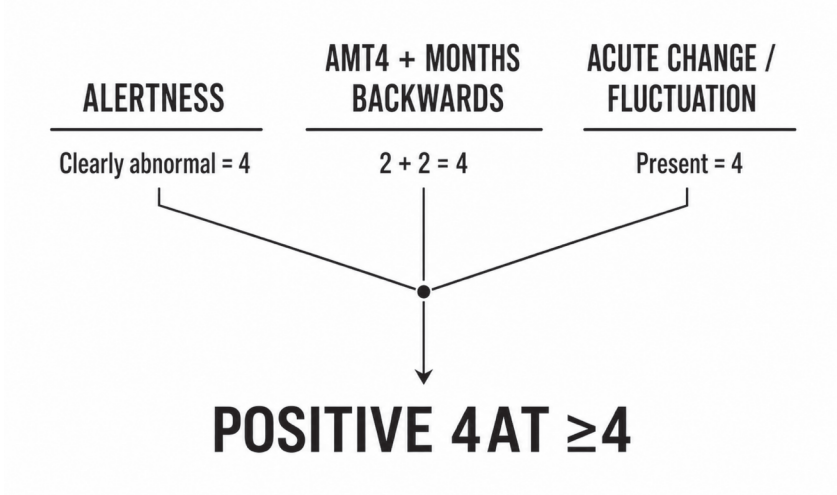


Figure 1. The three routes to a positive 4AT score of 4 or more.

EARLY VERSIONS WERE TESTED REPEATEDLY in real clinical practice, revised in response to user feedback, and the finished tool was published on the website in 2011. The current version, 1.2, dates from 2014, with only minor wording changes since the original: stability is deliberate, because the validation evidence attaches to the published wording. The tool was designed by Alasdair MacLullich of the Edinburgh Delirium Research Group at the University of Edinburgh, with Tracy Ryan and Helen Cash of NHS Lothian, with generous help from colleagues named in the Acknowledgements.

The name records the design: four items, each beginning with A, forming an Assessment Test and, conveniently, a test whose job is to alert.

What happened next took several years. The first validation study, led by Giuseppe Bellelli in Italy, appeared in 2014 [9], and a steady accumulation of diagnostic accuracy studies followed across emergency departments, medical and surgical wards, stroke units, hospices and care settings: 33 studies in over 6,000 patients and some 20 languages, one of the largest formal validation bases for any delirium tool (Chapter 10). By 2021 a survey of UK acute healthcare organisations found the 4AT the most widely used delirium tool in the country [10], the randomised comparative diagnostic accuracy study against the CAM had reported [11], and in 2023 the UK's National Institute for Health and Care Excellence updated its delirium guideline, replacing the short Confusion Assessment Method (CAM) with the 4AT as the recommended tool outside critical care; the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) and Intensive Care Delirium Screening Checklist (ICDSC) continue to be recommended in critical care, on the grounds that it "is among the most accurate of the tools reviewed, quick and simple to use, and has a broader range of evidence to support it" [12]. The 4AT now appears in guidelines and care standards in several countries and is in routine use in health systems worldwide. Its total global footprint is not measurable, since no international registry of delirium tool use exists, but adoption has been rapid; it is the most used delirium tool in UK acute hospitals [10], and among general

(non-ICU) delirium instruments it is one of the most widely implemented. Inside intensive care, the CAM-ICU remains the most widely used instrument internationally.

For all that, the 4AT remains what it was designed to be: a two-minute bedside test that any clinician can do.

DELIRIUM: THE ESSENTIALS

A woman is admitted with pneumonia. Three days later she is quiet, sleepy, and looks through her daughter rather than at her. The night staff say she was trying to climb out of bed at 4am, convinced the ward was a ferry terminal. By morning she is subdued again. This is a typical presentation of delirium.

This chapter gives the minimum knowledge needed to use the 4AT well. For fuller coverage, see the delirium care guide at www.the4AT.com/deliriumguide and the further reading in Chapter 15.

WHAT DELIRIUM IS

Delirium (from the Latin *delirare*, to leave the furrow) is an acute disturbance of attention, awareness and cognition, caused by one or more stressors: acute illness, surgery, drugs, drug withdrawal, physiological disturbance, and sometimes the psychological stress of illness and admission itself [13,14]. The definition is essentially the same in DSM-5-TR (the fifth edition, text revision, of the Diagnostic and Statistical Manual of Mental Disorders) and ICD-11 (the eleventh revision of the International Classification of Diseases).

By definition, all delirium is of acute onset. It develops over hours

to days, it tends to fluctuate, often worse in the evening and overnight, and the disturbance is always a change from the person's baseline. That last clause is the working core of the definition, and it is why Item 4 of the 4AT exists.

Delirium takes three broad motor forms. Hyperactive delirium, with restlessness, agitation and sometimes hallucinations or paranoid beliefs, is the form everyone recognises. Hypoactive delirium, with drowsiness, reduced movement and quiet inattention, is more common, more often missed, and carries at least as poor a prognosis: the "good patient" who causes no trouble and is sometimes dying quietly of an undetected problem. Mixed delirium moves between the two, sometimes within an hour. Hypoactive presentations have their own differential: depression and, rarely, catatonia can look similar at the bedside, and the acute time course, informant history and attention testing are what separate them. Some patients do not show clear motor disturbances either way.

Psychotic features are common and under-recognised: delusions, often paranoid, and hallucinations, usually visual. Many patients are frightened, and some are extremely frightened. A proportion remember the episode afterwards in distressing detail. It is worth remembering what detection actually amounts to: finding a seriously ill person in an acute and often terrifying brain failure, who cannot ask for help.

HOW COMMON, AND HOW SERIOUS

Delirium affects at least one in seven adults in general hospitals, and a larger proportion of older inpatients, with much higher rates again in some groups: after hip fracture and other major surgery, in intensive care, in advanced illness, and in people with dementia [13,15].

The outcomes are serious. Delirium is associated with higher mortality, longer stays, more institutionalisation, and higher subsequent rates of dementia [13]. In a two-centre study of 82,770 emergency admissions in Scotland, a positive 4AT at admission was associated with higher mortality, longer stays and less time at home

after discharge [16]. In 18,040 hip fracture patients across the UK, delirium detected by the 4AT at presentation was associated with death, longer stays and loss of independence [17]. Whatever fraction of these outcomes delirium causes rather than marks, the practical conclusion is the same: a positive test identifies a patient at high risk who needs attention now.

Delirium is also, in a meaningful proportion of cases, preventable. Multicomponent prevention programmes addressing orientation, sleep, mobility, hydration, nutrition, vision and hearing reduce delirium incidence in hospital populations [18]. Prevention is beyond this manual's scope, but a detection programme sits naturally beside one.

THE DETECTION GAP

Many studies show that delirium is often missed. Some institutions miss more than 80% of cases, and poor detection unfortunately remains the norm in many places. There are exceptions in which systematic assessment has been achieved at scale: in a national hip-fracture study, 4AT completion reached 91% [17]. The reasons for missed cases are familiar: hypoactive presentations that look like fatigue; fluctuation, so the patient seems fine on the ward round; the word “confused”, which sounds like a description but usually stops anyone looking further; the assumption that an older person with dementia is “always like this”; and, historically, tools too long or too training-hungry for routine use.

The word delirium stimulates action in a way that “confusion” does not. It signals an emergency, triggers a search for causes, warns of risk, and gives patient and family an explanation and a name. Nothing else in delirium care can happen until the delirium has been detected [12,15].

DELIRIUM AND DEMENTIA

The two conditions are entangled in every direction. Dementia is the largest risk factor for delirium; delirium is common in people with dementia (delirium superimposed on dementia); an episode of delirium is associated with accelerated cognitive decline and new dementia diagnoses; and at a single bedside encounter the cognitive picture of the two can be identical. Time course separates them. Delirium arrives in hours to days and fluctuates, whereas dementia arrives over months to years and is comparatively stable, which is why informant history matters so much, why Item 4 carries 4 points, and why the interpretation of untestable scores in severe dementia needs care (Chapter 4).

A note the other way round: a hospital episode with a positive 4AT is also a signal for possible underlying dementia. In a study of 75,221 admissions, higher 4AT scores were strongly associated with an existing or subsequent dementia diagnosis [19]. So a delirium detection programme has a second benefit: it also picks up people who may have undiagnosed dementia.

WHAT THE 4AT ASKS OF YOU

Using the 4AT requires no course, no certificate and no psychometric background. It does require the ten minutes of background in this chapter, the item rules in the next one, and a willingness to take a two-minute test seriously. It is a short assessment but it is still a clinical assessment. The patients most likely to score positive are usually in no position to say anything if it is done carelessly.

ADMINISTERING THE 4AT: ITEM BY ITEM

Before starting, do the usual steps that optimise assessment as much as possible. Make sure hearing aids are in and switched on, and glasses are on. Reduce background noise if you can. Face the patient, speak clearly, and allow time for answers. If the patient's first language is not yours, involve an interpreter where possible (see the end of this chapter).

Introduce yourself and explain that you are going to ask a few routine questions. Most patients are familiar with this. The bedside part of the assessment usually takes around a minute.

ITEM 1: ALERTNESS

What you are assessing. The patient's level of alertness during your assessment, rated by observation. Abnormal alertness runs in two directions: reduced (drowsy, difficult to rouse, drifting off during the assessment) and increased (agitated, hyperactive, hypervigilant). Both score 4.

How to do it. Observe from the moment you approach. If the patient is asleep, attempt to wake them with speech, or a gentle touch on the shoulder. Ask the patient to state their name and address: this

standard opening gives you something to rate. Keep observing throughout the rest of the test, because alertness that deteriorates or fluctuates during the assessment counts.

SCORING

Score 0 — Normal: fully alert, not agitated, throughout

Score 0 — Mild sleepiness for less than 10 seconds after waking, then normal

Score 4 — Clearly abnormal

The judgement in practice. Plenty of hospital patients are sleepy: it is 3pm, they were clerked at midnight, the ward is warm. A patient who is briefly heavy-eyed on waking but then converses normally scores 0. The abnormality you are looking for is qualitatively different: the patient who needs repeated stimulation to stay with you, whose eyes close mid-sentence, who gives you a few words and fades. At the other end, restlessness beyond ordinary discomfort, hypervigilance, or obvious agitation scores 4. If you are having to work to keep the patient engaged, that in itself is what you are looking for. Score Item 1 as abnormal.

Why does this single observation carry 4 points? In acute general hospital populations, once obvious alternative explanations are set aside (deep sedation, coma, known severe neurological injury), the great majority of patients with clearly altered alertness prove to have delirium, and the published arousal literature points the same way [1–3,20]. It is also the item that catches hypoactive delirium, the form most often missed.

Pitfalls. Recent sedatives, opioids or a post-anaesthetic state can depress alertness: score what you see, and let the clinical assessment that follows sort out cause (drug-induced altered alertness is itself frequently part of a delirium). Do not score deafness or dysphasia as drowsiness: a hearing-impaired patient who is awake and attending scores 0 here regardless of how the conversation goes. And do not

talk yourself out of the finding because the patient perks up briefly. Delirium fluctuates; a short lucid patch is expected.

ITEM 2: AMT4

What you are assessing. Orientation and basic memory, using four questions from the Abbreviated Mental Test [4,5]: age, date of birth, place (the name of the hospital or building), and the current year.

How to do it. Ask directly, one at a time: “How old are you?”, “What is your date of birth?”, “Where are we just now: what is this place called?”, “What year is it?” Do not give hints. Do not react to errors; move to the next question. Accept a self-correction if it comes without prompting.

SCORING

Score 0 — No mistakes

Score 1 — 1 mistake

Score 2 — 2 or more mistakes, or untestable

What counts as correct. For place, any reasonable identification of the setting is acceptable: the hospital’s name, “the hospital”, “the infirmary”, “A&E”. The full formal title is not required. For year, the exact year is required; in the first days of January, sensible judgement applies to a patient answering with the year just ended. For age and date of birth, the answers should match the records.

Untestable. If the patient cannot make a meaningful attempt because of drowsiness, severe inattention or inability to produce speech, score 2 and record why. Untestable is a rating in its own right and it contributes to the score (see Chapter 4). Do not score a patient untestable merely because of language: use an interpreter, family member, or written questions where appropriate, and see the section on communication below.

Pitfalls. Confident wrong answers are still wrong: confabulation

is common in delirium and in dementia. A patient with dementia may know age and date of birth (long-stored facts) but miss place and year (current orientation); a patient with delirium often misses across the board. The pattern is clinically informative even though the score does not use it.

ITEM 3: ATTENTION (MONTHS BACKWARDS)

What you are assessing. Sustained attention, the core cognitive deficit of delirium, using Months of the Year Backwards, one of the best-evidenced brief attention tests [6,7,21].

How to do it. Say: “Please tell me the months of the year in backwards order, starting at December.” To assist initial understanding, one prompt is permitted: “What is the month before December?” Then let the patient proceed without further help. Allow thinking time; slow but accurate performance is acceptable. Self-corrections are acceptable. Stop when the patient has clearly succeeded (seven or more months correct, that is, December back to June), or has made an uncorrected error, or cannot continue.

SCORING

Score 0 — Achieves 7 months or more correctly

Score 1 — Starts but scores fewer than 7 months, or refuses to start

Score 2 — Untestable (cannot start because unwell, drowsy, inattentive)

The boundary between 1 and 2. This trips people up. If the patient engages with the task at all, even just “December...” and no further, or verbally declines, score 1. Score 2 only if the patient cannot start because of drowsiness or severe inattention. The distinction matters because untestable status on both cognitive items adds to 4 points, flagging possible delirium, which is exactly what severe inattention should do.

Pitfalls. Some patients begin reciting months forwards, or list seasons; restate the instruction once, and if the misunderstanding persists, score 1. Skipped months are errors. Omitting the test because the patient “seems fine” wastes the item most sensitive to early or mild delirium; patients with fluent social conversation can fail months backwards impressively. Education and language affect performance, so weigh a borderline result accordingly, and remember the test can be done in the patient’s own language. What matters is whether they can sequence the months, not which language they do it in.

ITEM 4: ACUTE CHANGE OR FLUCTUATING COURSE

What you are assessing. Evidence of significant change or fluctuation in alertness, cognition, or other mental function (for example new paranoia or hallucinations), arising over the last two weeks and still evident in the last 24 hours. This is the feature that separates delirium from dementia and from lifelong impairment, and it is present in all diagnostic definitions of delirium [14].

How to do it. Use every source available: your own knowledge of the patient, ward and care home staff at every grade (the healthcare assistant who fed the patient yesterday often knows more than anyone), family and carers, the GP letter, previous notes and discharge summaries, therapy and functional assessments, and prior cognitive test results including earlier 4ATs. Useful questions to an informant: “Has there been a sudden change in her thinking or alertness in the last couple of weeks?”; “Does he come and go during the day: drowsy and muddled at times, clearer at others?” Two questions to the patient can surface psychotic features politely: “Are you concerned about anything going on here?” and “Have you been seeing or hearing anything unusual?”

SCORING

Score 0 — No

Score 4 — Yes

When there is no informant. At the front door, a quarter or more of older patients arrive with no one who can describe their baseline. The 4AT was deliberately designed so that this does not stall the assessment. First, an informant can often be reached by phone shortly afterwards. Second, the clinical course itself can establish acute change: a patient who is drowsy and incoherent, but who the notes show was living independently ten days ago, has changed, and Item 4 can be scored 4 on that evidence; this is a reasonable clinical judgement. Third, even with Item 4 scored 0 for lack of information, the 4AT can still reach 4 or more through Item 1 or through untestable ratings on Items 2 and 3. If a tool cannot produce a result until an informant has been found, the diagnosis of a medical emergency waits on a phone call. The 4AT was designed to avoid that.

Weighing informant quality. Not all collateral history is reliable. A relative who visits twice a year may not know the baseline; occasionally the informant has cognitive impairment themselves. Treat informant history like any other clinical evidence: seek it always, weigh it in context.

A note on fluctuation. Some fluctuation occurs in dementia without delirium, particularly dementia with Lewy bodies. Marked fluctuation, though, usually means delirium. If the picture is dominated by fluctuating cognition in someone with suspected or known Lewy body disease, the assessment needs senior clinical input rather than a scoring rule.

Item 4 incorporates the logic of the Single Question in Delirium (SQiD), which asks someone who knows the patient: “Do you think [patient’s name] has been more confused lately?” [22]. Variants are widely used, including versions asking whether the person is more confused or drowsy than usual

(Chapter 11). If a SQiD has already been asked, its answer feeds directly into this item.

COMMUNICATION DIFFICULTIES: DEAFNESS, DYSPHASIA, LANGUAGE

The 4AT is more adaptable than tools that depend entirely on a verbal interview with the patient, because two of its four items (1 and 4) require no conversation at all. For the cognitive items, adapt sensibly and record what you did. For deaf patients, write the questions down; the AMT4 and Months Backwards both work on paper. For patients with dysphasia after stroke, interpretation is genuinely difficult: involve the speech and language therapist, leaning on Items 1 and 4, and treat cognitive item scores with caution. For patients without English, use an interpreter (professional where possible), ask the AMT4 in translation, and do Months Backwards in the patient's own language; validated translations of the whole tool exist in more than 20 languages (Chapter 12). Score untestable only when the barrier is the patient's clinical state, not the language; where a language barrier genuinely prevents testing and no interpreter is available, record that the cognitive items could not be administered and rely on Items 1 and 4, which usually remain scoreable.

SCORING AND INTERPRETATION

THE SCORE

Add the four item scores. The total runs from 0 to 12.

4 or above — Possible delirium, with or without cognitive impairment

1 to 3 — Possible cognitive impairment

0 — Delirium or severe cognitive impairment unlikely

Three qualifications go with those bands.

A score of 4 or above is not a diagnosis of delirium; it is a strong signal that requires clinical assessment (Chapter 5). A score of 1 to 3 does not exclude delirium; it usually reflects cognitive impairment without current evidence of acuity, and it calls for more detailed testing and informant history in due course. A score of 0 makes delirium unlikely but cannot fully exclude it, particularly when the Item 4 information was thin; if suspicion persists, reassess, and ask someone who knows the patient.

THE ROUTES TO A POSITIVE SCORE

How a patient reaches 4 matters as much as whether they do, because the routes mean different things, and reading them is most of the interpretive skill the 4AT asks of its user.

Item 1 alone (abnormal alertness). A highly specific delirium signal in acute settings. Treat as possible delirium regardless of anything else.

Item 4 alone (acute change or fluctuation). Also highly specific. A patient with normal alertness and perfect cognitive testing can still score 4 this way, and rightly: delirium fluctuates, and a lucid interval at the moment of testing proves nothing. New paranoid beliefs or hallucinations reported over recent days count here.

Items 2 and 3 both untestable (2 + 2). The patient who cannot engage with either cognitive item scores 4. In acute presentations this usually indicates delirium. The main exception is severe dementia, discussed below.

Combinations. Cognitive errors plus acute change (for example 1 + 1 + 4) is a common pattern in delirium superimposed on early cognitive impairment. High totals (8 to 12) are associated with an even higher probability of delirium.

READING THE PATTERN AS WELL AS THE TOTAL

The total puts the patient into one of the three bands. The pattern of item scores tells you what is going on. A 6 made of abnormal alertness and two minor cognitive errors, with no evidence available for Item 4 (4 + 1 + 1 + 0), is a different patient from a 6 made of normal alertness, two minor cognitive errors and overnight fluctuation (0 + 1 + 1 + 4). Document all four item scores with the total, and consider writing in a few words why each item scored: “Item 1: repeatedly drifting off during questions”. The next assessor, comparing tomorrow’s 4AT with yours, will thank you, and repeated 4ATs are much

more useful if the item scores are recorded each time and not just the total (see recovery monitoring, Chapter 5, and the documentation example in Chapter 14).

SEVERE DEMENTIA

People with severe dementia may be unable to answer the AMT4 or attempt Months Backwards at their baseline, with normal alertness and no acute change. They will score $2 + 2 = 4$: possible delirium, by the tool's rules, without delirium being present. This is a known and accepted trade-off made at the design stage. Treating untestable patients as unscorable would leave without a result the group most likely to have delirium, in order to avoid flagging a smaller group who have severe dementia at baseline. A missed delirium in an acutely unwell patient carries more risk than an unnecessary clinical assessment in a patient with dementia. What it means in practice: the 4AT's specificity is slightly lower in severe dementia than in general populations [23,24], and a score of 4 in someone with known severe dementia is a prompt to do exactly what a careful clinician would do anyway: establish the baseline from someone who knows the person, and look for change. If the care home staff say "this is her normal", with normal alertness and nothing new, the likely conclusion is severe dementia without delirium, with continued monitoring for change. If nobody knows the baseline, assume possible delirium and assess. If you are going to be wrong, that is the better way to be wrong.

Deciding whether someone has severe dementia is itself unreliable during a possible delirium episode, so the baseline question ("what was she like two weeks ago?") is more useful than anything observed at the bedside. In this group, diagnosis depends to a large extent on informant history and observation over time.

A simple combined arousal and attention testing procedure has been evaluated specifically for detecting delirium superimposed on dementia [25]. A separate instrument developed for the same prob-

lem, the 4-DSD (a four-item tool for delirium superimposed on dementia), is described in Chapter 13.

BORDERLINE SCORES

A score of 3 usually means cognitive errors without evidence of acute change, but it can also be early or fluctuating delirium caught at a better moment, especially when the Item 4 information was incomplete. For patients scoring 1 to 3 in acute settings: treat as at high risk of delirium, consider baseline cognitive work-up in due course (Chapter 5), and have a low threshold for repeating the 4AT if anything changes, or simply later the same day if suspicion nags.

FACTORS THAT AFFECT THE SCORE

Drugs. Sedatives, opioids and anticholinergic drugs can depress alertness and blunt cognition. This is usually not a false positive in any useful sense: drug-induced disturbance of arousal and cognition in an unwell older person commonly is delirium, or is about to be. Score what you find; review the chart as part of the response.

Severe pain, breathlessness, and the moment of testing. A patient in crisis performs badly. Where possible, test when reversible discomfort has been addressed; where not possible, score what you find and note the context.

Education, language and culture. The AMT4 and Months Backwards work reasonably well across cultures, but they are not culture-free. Date of birth conventions differ; months may be more natural in another calendar or language; low formal education affects backwards sequencing. Adapt delivery (Chapter 3), note any modification, and weigh borderline cognitive-item scores with this in mind.

Immediately after anaesthesia. Scores in the first minutes after anaesthesia partly reflect emergence rather than delirium; the post-anaesthesia validation work waited until recovery-room criteria were met before testing [26]. Time the assessment accordingly, then take a

positive score seriously: delirium in the recovery room is associated with worse outcomes.

REPEATING THE 4AT

The 4AT is an episodic test, and single snapshots may miss fluctuations in mental state. Repeat it when the clinical picture changes, when staff or family report new confusion, after a positive monitoring signal (Chapter 7), daily for several days after major surgery in older patients, and daily to track recovery from a diagnosed episode (Chapter 5). Repeated cognitive testing several times a day for days on end, though, burdens patients and staff and invites practice effects; that is monitoring's job, done with different instruments (Chapter 7).

WHAT TO DO AFTER A POSITIVE 4AT

This chapter is about what a positive score of 4 or more should set in motion. This is an overview, not a treatment textbook: for fuller delirium management guidance, see the Scottish Intercollegiate Guidelines Network (SIGN) and NICE guidelines [12,15] and the care guide at www.the4AT.com/deliriumguide.

SCORE 4 OR ABOVE: POSSIBLE DELIRIUM

First, assess, and establish whether delirium is present. The score's job is to trigger a prompt clinical assessment: history including baseline and time course, examination, drug review, and targeted investigations. The diagnosis of delirium is made clinically, against standard criteria [14], by the assessing clinician. Sometimes the conclusion will be that delirium is not present, for instance because the presentation reflects baseline severe dementia. That is a normal and expected outcome of this assessment process, not a fault in it. If delirium is confirmed, proceed.

Search for causes, and expect there to be more than one. Delirium in older people usually has several contributing causes: infection, yes, but also drugs (opioids, benzodiazepines, anticholiner-

gics, recent changes and recent withdrawals), dehydration and electrolyte disturbance, hypoxia, urinary retention, constipation, pain, organ failure, and alcohol withdrawal. The commonest mistake in practice is to find one plausible cause, often a urine dipstick, and stop looking.

Treat the causes, and care for the brain in the meantime. Reverse what is reversible. Alongside that, provide supportive care: reorientation and familiar faces, glasses and hearing aids, day-night routine and sleep protection, early mobilisation, hydration and nutrition, avoidance of unnecessary catheters, moves and restraints, and regular review. Most delirium can be managed this way by generalist teams on ordinary wards; it does not need a specialist to do it well. Antipsychotics and benzodiazepines have no role in treating delirium itself; their only place is severe distress or psychosis unrelieved by non-drug measures, or immediate safety, at the lowest effective dose, for the shortest possible time. Exceptions: alcohol or benzodiazepine withdrawal or other specific indications. Antipsychotics should generally not be used in Parkinson's disease or Lewy body dementia.

REDUCED AROUSAL

A significantly reduced level of arousal matters immediately because it increases aspiration risk. Decisions about oral food, fluids and medicines should take account of the patient's level of alertness and swallowing safety.

Assess and document distress and psychotic features. Detection should not stop at "4AT positive". Ask whether the patient is frightened, paranoid, or hallucinating. These features guide treatment and matter enormously to the person; they are also the features families most need explained. Ask the two questions from Chapter 3, and write down what you find.

Tell the patient and family what is happening, always using the word 'delirium'. A named diagnosis, a simple explanation ("an acute disturbance of brain function caused by illness or drugs; common; usually improves as the cause is treated; not the same as dementia"),

and a leaflet or a pointer to the family guide at www.deliriumsupport.com change the experience of one of the most frightening events a family will witness. A delirium episode also matters for the patient's future: document it in the discharge summary, because a history of delirium predicts recurrence and flags dementia risk [19].

Document the diagnosis as well as the score. An entry reading “4AT 6/12; delirium confirmed on assessment; likely causes X and Y” tells the next person far more than “4AT 6” alone.

SCORE 1 TO 3: POSSIBLE COGNITIVE IMPAIRMENT

No delirium is evident now, but the score does not exclude it. In an acute setting this patient is at elevated risk of developing delirium, so preventive care and a low threshold for reassessment apply. The findings also deserve consideration of follow-up in their own right: many dementias first surface as an “incidental” cognitive score in hospital. Acute illness is the wrong moment for formal cognitive testing for dementia. Approaches vary, but it is important not to ignore detected cognitive impairment. Some hospitals have a process of inpatient referral for possible dementia if the history suggests it (cognitive and functional decline over several months or years). Others use the route of documentation, information to the GP, and cognitive assessment once the acute episode has settled, typically after six to twelve weeks, with informant history (and, where available, tools such as the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) or the AD8 (Eight-item Informant Interview to Differentiate Aging and Dementia)) meanwhile. Do not label anyone with dementia on the strength of a hospital cognitive score alone.

SCORE 0

Delirium and moderate to severe cognitive impairment are unlikely. Record the result: a documented 0 is valuable, because it timestamps a normal baseline against which any later change can be judged.

If clinical suspicion is strong despite the o (a family member insisting “he’s just not right”), trust the suspicion enough to reassess later and to seek more informant history. Because delirium fluctuates, one normal assessment does not always exclude it. Good practice is to remain alert to fluctuation and repeat the assessment if a change is apparent.

TRACKING RECOVERY

Delirium can resolve quickly (over a small number of days) or can end slowly and unevenly, and knowing where a patient is on that path informs discharge decisions, rehabilitation and family conversations. Repeating the 4AT, typically daily during and after an episode, gives a simple, validated way to follow recovery: a study by Haruno McCartney, Erin Noble and colleagues published in 2025 formally validated the 4AT for assessing recovery from delirium in older hospital patients [27]. Falling scores usually track genuine improvement; a stalled or rising score is a prompt to look again for unresolved causes. Item-level documentation (Chapter 4) makes these trajectories far more informative: alertness recovering while orientation lags tells a different story from the reverse.

Two things should happen before the episode is treated as closed. First, before discharge, check for and document any persistent cognitive impairment, arrange follow-up, and ensure the delirium episode is coded and communicated. Second, offer the patient the chance to talk about what they experienced; a proportion carry vivid, frightening memories, and a short conversation, with written information, is a simple and humane intervention.

THE 4AT ACROSS SETTINGS

The items themselves are the same wherever the 4AT is used. What varies between settings is when it is done, who is available to give a history, and what happens after a positive score. This chapter takes the settings in turn. Broadly, the 4AT is intended to be used at these moments in care: at presentation to acute settings; before and after surgery; at transitions between services; whenever delirium is suspected; after a positive monitoring signal; and daily to track recovery during an episode. However, some units have used the 4AT differently, as a daily or twice-daily monitoring tool, sometimes substituting Days of the Week Backwards, which is less burdensome when the test is repeated often.

EMERGENCY DEPARTMENTS AND ACUTE FRONT DOORS

Most delirium in a hospital's day arrives through the front door, so this is where systematic assessment has a high clinical yield, enabling earlier intervention. The evidence base here is broad: validation studies in emergency department (ED) and acute medical populations include the multicentre UK trial [11], Irish work in older ED

attendees [28], the French-Canadian ED study [29], Spanish ED work including use at triage [30,31], and Norwegian experience in suspected sepsis [32].

Practicalities. Assess older patients (65 or 70 and over by local policy) at presentation, alongside vital signs, and any adult in whom delirium is suspected. The two-minute run time is compatible with triage; one Spanish study embedded it there [31]. Informants are often absent at the front door: use the ambulance sheet, the GP letter, phone the care home or family, and remember the routes to a score that need no informant (Chapter 4). A positive score in the ED needs to be handed over explicitly. Delirium at presentation is associated with mortality and long stays [16,17], and the receiving team should be told about it rather than left to find it on the form.

For patients being discharged from the ED with a positive score, someone must own the follow-up: delirium at home needs a cause found, a carer informed, and early review arranged.

ACUTE MEDICAL AND GERIATRIC WARDS

On the wards, the 4AT has two distinct jobs, and separating them is helpful in designing implementation.

The first is assessment at admission: every older patient within, say, 24 hours of arrival, establishing presence or absence of delirium and a documented baseline. The second is assessment when something changes: a positive monitoring signal (Chapter 7), a nurse's concern, a family comment, a fall, new drowsiness. Norwegian quality improvement work found nurses could run admission assessment at scale, with completion rates near 90% [33].

In the author's view, doing a 4AT on every patient every day indefinitely is not the best use of the tool, because of the burden on patients and staff. However, some units use the 4AT for high risk parts of the patient stay and reduce the burden by changing to Days of the Week Backwards as mentioned above.

SURGERY: BEFORE AND AFTER

Postoperative delirium is common in older surgical patients, with risk varying by operation and baseline factors, and it is routinely missed without assessment. The 4AT fits three perioperative moments. Before surgery: a baseline 4AT identifies existing impairment and delirium, stratifies risk, and gives the post-op assessments something to be compared against. In the recovery room: German work validated the 4AT in the post-anaesthesia care unit, testing once recovery criteria were met [26]. On the ward: daily assessment for the first three to seven days, when incidence peaks, is a reasonable default in older or high-risk patients, with cardiac surgical validation supporting use in that population too [34].

Post-anaesthesia timing matters: score once the patient has met recovery-room criteria, not during emergence, and interpret early drowsiness with the anaesthetic in mind (Chapter 4).

HIP FRACTURE AND ORTHOGERIATRICS

Hip fracture patients have a high risk of delirium because of multiple potential risk factors, including older age, frailty, dementia, pain, opioids, surgery. National audits in the UK have made 4AT assessment a standard of care in this pathway, and the results form the largest published dataset on how a delirium tool performs when embedded in a national pathway: across 18,040 patients in the national hip fracture registry work led by Rose Penfold, 4AT completion was 91%, 21% scored positive, and positive scores were associated with death, longer stays and new residential care [17]. During the COVID-19 period, the same pathway data showed delirium marking out patients with much worse outcomes [35]. Practice points: assess at presentation, before surgery where feasible, and daily post-operatively; treat a positive pre-operative 4AT as a flag for the whole perioperative team, including anaesthetics.

INTENSIVE CARE

The 4AT was designed for general settings, not for intubated patients. In ventilated or sedated patients, the intensive care unit (ICU)-specific tools, the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) [36] and the Intensive Care Delirium Screening Checklist (ICDSC) [37], remain the standard recommended instruments. The 4AT has a place at the edges of intensive care: in patients who can speak, at step-down, and at the ICU-to-ward transition. A four-centre Korean study in critically ill (non-ventilated) patients found the 4AT performed acceptably against a DSM-5 reference, sensitivity 74% and specificity 95% [38], with the practical advantage that ward teams already know it. In practice the division of labour is clear. Use the ICU's own tools while the patient is in the unit, and the 4AT once the patient is extubated and potentially able to hold a conversation and at handover to the ward.

PALLIATIVE CARE AND HOSPICES

Delirium affects at least 1 in 3 people who are in the last few days of life, with the prevalence higher as death approaches. Detecting delirium really matters because it can help improve the management of distress. A two-hospice Scottish validation study led by Elizabeth Arnold found the 4AT performed strongly in hospice inpatients (sensitivity 89%, specificity 94% against a DSM-5 reference) [39], and earlier hospice quality-improvement work by Baird and Spiller had already shown it usable on admission [40]. In the context of palliative care, informants are usually present and know the baseline well, making Item 4 unusually reliable; hypoactive presentations dominate and are easily misread as dying, fatigue or opioid effect; and the purpose of assessment shifts with the goals of care, from finding reversible causes early on, towards comfort, explanation to family, and management of distress when death is near. The 4AT itself is used in the same way as anywhere else. What changes is how much is done in response to a positive score.

CARE HOMES

Care home residents have higher rates of delirium than people living at home. This is because they are an old and frail population with very high rates of dementia (more than 70%). Staff know each resident's baseline far better than hospital staff ever could. That baseline knowledge makes Item 4 powerful in this setting, and it makes care home staff ideal SQiD reporters ("is she more confused or drowsy than usual?") feeding a 4AT done by visiting clinicians. Use the 4AT when a resident changes: new confusion, new drowsiness, a fall, reduced eating and drinking, "just not herself". A positive score is the start of the same search for causes as anywhere else, and often the deciding information in a conveyance decision. The severe dementia caveat (Chapter 4) applies particularly commonly here; the staff's baseline report is the anchor.

COMMUNITY, PRIMARY CARE AND AMBULANCE SERVICES

GPs, community nurses, paramedics and hospital-at-home teams meet delirium first, usually as a story: found wandering, stopped eating, "gone downhill since Tuesday". The 4AT needs no equipment and works at a kitchen table. For paramedics, a documented 4AT together with a carer's account of when things changed is extremely useful at handover. A written score gets passed on; a verbal impression usually does not. For primary care, the 1 to 3 band has a particular after-use: once the acute episode has resolved, it points to the patients who may need memory assessment (Chapter 5). In community settings the tool itself is rarely the difficulty. What needs agreeing locally, in advance, is the pathway: who gets phoned, who reviews the patient, and when admission is the right call.

STROKE UNITS AND OTHER SPECIALIST SETTINGS

Delirium complicates around a quarter of acute strokes, and assessment must contend with dysphasia; stroke-unit studies have applied the 4AT with the adaptations discussed in Chapter 3, and Norwegian nurses ran it as routine assessment in a stroke unit at scale [42–44]. In Parkinson’s disease and Lewy body disease, fluctuation and hallucinations occur without delirium, and a 2025 study by Rachael Lawson and colleagues examined bedside tools including the 4AT in Parkinson’s inpatients [45]: scores in this group need interpreting with the disease in mind and with senior input. Validation also extends to traumatic brain injury via the Traditional Chinese version [46] and to skilled nursing facilities (Chapter 10).

DETECTION AS A SYSTEM: EPISODIC TESTS AND MONITORING

More than 60 delirium detection instruments have been published (see www.deliriumtools.com, maintained by the author). Faced with that number, services usually ask which tool is best. It is more useful to work out first what job the tool has to do, because the answer differs. There are two different jobs in delirium detection, and they need two different kinds of instrument.

The first job is the episodic assessment: a definitive-as-possible bedside answer to “does this patient have possible delirium right now?” This is the job at the front door, perioperatively, at transitions, and whenever suspicion arises, and it demands a balance of sensitivity and specificity good enough to act on, which in practice means built-in cognitive testing and a change-from-baseline judgement. The 4AT is designed as an episodic tool. So is the CAM, done properly (as specified in the CAM manual, with a cognitive test and interview done before the CAM is scored). Episodic tools require bedside assessment and interaction with the patient, and most involve some interview questions and/or cognitive testing. This is entirely feasible when done a limited number of times, but can be burdensome if done daily (especially multiple times per day) for several days at a

time. In the author's experience a couple of minutes of bedside testing for a few days in a row is entirely reasonable, but can feel excessive if done twenty times a week per bed for all patients through the whole of their stay.

The second job is monitoring: watching for new delirium arising in patients already admitted, day after day. Monitoring has to be light enough that it actually gets done several times a day, for weeks, without wearing out staff or patients. In practice that means something that takes a few seconds and is mostly observational, because no patient should face formal cognitive testing three times a day for a fortnight (practice effects corrupt the results anyway). Monitoring instruments include the Single Question in Delirium (SQiD), a single question to someone who knows the patient ("Do you think [patient's name] has been more confused lately?") [22], and observational scales; in the UK, NEWS2 [47] includes new confusion or delirium within its ACVPU consciousness item, so routine observations already carry a delirium signal. Many NHS services pair that signal with the SQiD, with a positive result from either triggering a 4AT. This is a widely adopted model rather than a nationally specified process. Other monitoring instruments include the Nursing Delirium Screening Scale (Nu-DESC) [48] and the Delirium Observation Screening Scale (DOSS) [49]. Unfortunately, this remains an element of delirium research and practice which is not yet settled. In practice, there are multiple models in use [41]. Published experience with unofficial, shortened CAM use in electronic records, with cognitive testing omitted, is an example of a widely-implemented but ineffective monitoring system: high completion rates but positive-score rates far below the expected prevalence of delirium, suggesting that many, and perhaps most, cases are missed while documentation looks complete [50]; see also Chapter 11.

The system ideally should be a loop. Episodic 4AT at entry and at need; light monitoring in between; any positive monitoring signal answered with a 4AT; a positive 4AT answered with clinical assessment; and daily 4ATs to track an episode to recovery. Used this way, each instrument is being asked to do only what it was designed for.

Two important issues to consider. First, using an episodic tool as a monitoring tool tends to fail quietly rather than visibly: completion collapses or the test decays into a tick-box with poor performance. This is the case with the CAM, as mentioned above. Second, using a monitoring tool as the definitive assessment fails the other way: The Single Question in Delirium (SQiD) alone is not adequate as a stand-alone delirium assessment tool. Neither type of instrument substitutes for the other.

As stated above, some services have actually adapted the 4AT for short-period repeated use. This is an evolving part of 4AT use and is being evaluated. Currently the 4AT's published accuracy relates to its use as an episodic tool (Chapters 10 and 12).

CLINICAL CASES

The 30 cases here are drawn from clinical practice across several settings, with details altered. They are teaching cases rather than a prevalence sample, so positive and difficult presentations are deliberately over-represented. They are chosen so that each teaches something different, and several show the awkward situations that scoring rules exist to settle: the patient who cannot speak, the resident with severe dementia, the lucid interval, the score that is probably a false positive. Some cases are adapted from the case examples published at www.the4AT.com; multiple further scored cases are available via that page.

Scores are given item by item, as they should be documented in practice. You are welcome to use these cases in teaching.

EMERGENCY DEPARTMENT

Case 1: Too drowsy to speak. No patient is Unable to Assess

Mrs A, 84, lives alone, with carers twice a day; she has dementia but usually manages at home and converses normally. She is brought in after

her daughter found her very sleepy and hot. The emergency department diagnoses pneumonia. On approach her eyes are closed; she opens them briefly when asked her name, attempts a word or two, incomprehensible, and does not respond to requests to attempt the AMT4 or Months Backwards. Her daughter is clear: this is nothing like her usual self.

Item 1, Alertness: 4 (clearly abnormal). **Item 2, AMT4: 2** (untestable). **Item 3, Attention: 2** (untestable). **Item 4, Acute change: 4** (clear change from baseline). **Total: 12/12.**

Comment. On an interview-based tool this patient risks being recorded as Unable to Assess, which is to say, left without a diagnosis. On the 4AT her inability to engage is itself scored as an informative clinical feature, and the total reflects multiple features compatible with delirium and the need for urgent clinical assessment while causes are treated. This is one of the key design features of the 4AT.

Case 2: No common language. An interpreter, not an untestable rating

Mrs B, 78, was visiting family when she became confused overnight. She speaks Polish and little English. She is alert and calm. Her granddaughter interprets while a professional interpreter is arranged. Via interpretation: she gives her age and date of birth correctly, thinks the year is 1998, and cannot name the place. Asked for the months backwards in Polish, she starts confidently, reaches September, then loses the thread. Her granddaughter says she was sharp two days ago.

Item 1: 0. **Item 2: 2** (two errors). **Item 3: 1** (started, fewer than 7 months). **Item 4: 4.** **Total: 7/12.**

Comment. A language barrier does not make a patient clinically untestable. The cognitive items can usually be administered in the patient's preferred language, using a professional interpreter where possible. Items 1 and 4 do not require verbal testing of the patient, although Item 4 usually requires collateral information. Here, the granddaughter provides immediate communication support while a professional interpreter is arranged. Record how the test was delivered. Where possible, a professional interpreter is helpful when

obtaining a more detailed history and undertaking the fuller clinical assessment.

Case 3: The morning zero and the evening six

Mr C, 81, attends after a fall. At 10am his 4AT is 0: fully alert, AMT4 correct, Months Backwards fluent to May, and his daughter, phoned, reports no significant change in alertness, cognition or behaviour during the preceding two weeks and says he was himself during the last 24 hours. She adds only, "I wondered if he seemed a bit odd on Sunday", but cannot describe a definite mental status change. He is admitted for observation. At 9pm he is awake and calm but says with conviction that his car is being towed and asks to leave. The evening 4AT shows normal alertness, one AMT4 error, Months Backwards to October, and a clearly documented acute change.

Morning: 0/12. Evening: Item 1: 0. Item 2: 1. Item 3: 1. Item 4: 4. Total: 6/12.

Comment. Delirium fluctuates; single snapshots can miss it. The morning 0 was still valuable: it documented that no 4AT features were evident at 10am, making the evening change much easier to interpret. The daughter's vague concern lowered the threshold for repeating the test but, without evidence of a definite change, did not by itself make Item 4 positive. Repeating the test does not mean the first one was wrong. Repeat testing is simply how an episodic test has to be used in a condition that fluctuates.

Case 4: Agitation counts as abnormal alertness

Mr D, 56, is brought in tremulous, sweating and frightened, two days after stopping drinking. He is hypervigilant, startles at noises, and cannot settle for questions. He manages his age and date of birth, misses place and year, and is too agitated and inattentive to initiate the sequence at all. His flat-mate confirms he was his usual self three days ago.

Item 1: 4 (agitated, hypervigilant). Item 2: 2. Item 3: 2 (untestable through agitation). Item 4: 4. Total: 12/12.

Comment. Abnormal alertness runs in both directions; Item 1 catches the hyperactive end as well as the drowsy one. Possible alcohol-withdrawal delirium is a medical emergency requiring urgent, specific treatment. The temporal history makes withdrawal a leading cause, but it does not remove the need to assess for other contributors. Assessment must include immediate consideration of Wernicke's encephalopathy, and parenteral thiamine should be given without waiting for confirmation. Alcohol-related brain injury and head injury are also part of the differential in this group. The 4AT makes the possible delirium syndrome explicit and supports handover alongside the local alcohol-withdrawal assessment and treatment pathway.

ACUTE MEDICAL WARD

Case 5: The quiet patient

Mrs E, 88, admitted three days ago with a chest infection, was "no trouble at all" overnight. She is propped up, eyes closing, food untouched. She rouses to voice, answers her age slowly, then drifts; repeated stimulation is needed to keep her with you. AMT4: one answer then silence. When asked for Months Backwards, she cannot engage with the instruction and makes no response. The healthcare assistant who fed her yesterday says she was chatting then. A new opioid prescription is on the chart.

Item 1: 4. Item 2: 2 (untestable after one answer, meaningful engagement not possible). **Item 3: 2. Item 4: 4. Total: 12/12.**

Comment. Hypoactive delirium is common and easily missed because it may present as quiet deterioration. A quiet patient is less likely than an agitated patient to trigger urgent review. Here the findings prompted immediate clinical assessment, including observations, oxygenation and respiratory status, medication review, hydration, and a renewed search for infection and other contributors, and the conversation with her family changed as a result.

Case 6: Delirium on top of dementia

Mr F, 79, has Alzheimer's disease; he lives with his wife, recognises his family, and enjoys the radio. Over 48 hours he has become markedly worse: up all night, not recognising his own kitchen, asking his wife who she is. On the ward he is alert. AMT4: knows age and date of birth, misses place and year. Months Backwards: December, November, then stops. His wife is unequivocal that this is new.

Item 1: 0. Item 2: 2. Item 3: 1. Item 4: 4. Total: 7/12.

Comment. His cognitive item scores may partly reflect his baseline dementia and are difficult to interpret alone. The strongest evidence for possible delirium is Item 4: a clear, dated, informant-witnessed change. In delirium superimposed on dementia, collateral evidence of change from baseline is essential. The score cannot separate the contribution of delirium from that of dementia, and it does not have to: it identifies an acute syndrome that needs assessing.

Case 7: Watching an episode end

Mrs G, 76, developed delirium during sepsis from a urinary source. Day 2: 4AT 12 (markedly drowsy, 4; AMT4 untestable, 2; Months Backwards untestable, 2; clear acute change, 4). Day 5: alert (0), one AMT4 error (1), Months Backwards to August (1), fluctuation settling but still reported by night staff (4): total 6. Day 8: alert (0), one AMT4 error that her daughter confirms is usual for her (1), Months Backwards correctly to June (0), and no acute change or fluctuation evident during the previous 24 hours (0): total 1. She goes home on day 10 with follow-up arranged.

Comment. The 4AT has been evaluated for repeated assessment during delirium recovery [27]. Serial scores give the team and family a shared trajectory, but they should be interpreted alongside symptoms, function and clinical readiness for discharge. Documenting item scores makes the trend readable: in her case, alertness normalised before attention. A stalled or rising score should prompt review for persisting or new contributors, while recognising that fluctuation can still occur.

Case 8: Overnight confusion, morning composure

Mr H, 70, admitted three days ago with a chest infection, scored 4AT 0 on admission. The night staff report he was restless and disorientated overnight, wandering into the wrong bay. On the morning round he is alert and pleasant. AMT4: misses the name of the hospital. Months Backwards: reaches August, then errors.

Item 1: 0. Item 2: 1. Item 3: 1. Item 4: 4 (documented overnight change and fluctuation). **Total: 6/12.**

Comment. The night staff's report is collateral information, and it counts towards Item 4. A patient does not need to display delirium features during the doctor's brief encounter for possible delirium to be recognised. If you insist on seeing the features yourself, you will miss a lot of fluctuating delirium. The admission 0 strengthens the evidence that this is new. Handover documentation ("restless, disorientated, wrong bay, 0300") captures exactly the evidence needed for Item 4.

SURGICAL AND PERIOPERATIVE

Case 9: The recovery room

Mrs I, 74, undergoes elective hip arthroplasty under general anaesthesia. At the point at which recovery-room discharge criteria were being assessed, she remains markedly drowsy, her eyes closing between words. The AMT4 and Months Backwards cannot be attempted. Her 4AT before surgery that morning was 0.

Item 1: 4. Item 2: 2. Item 3: 2. Item 4: 4 (clear acute change from the documented pre-operative baseline). **Total: 12/12.**

Comment. Timing matters: 4AT scores may be positive during immediate emergence but resolve a few minutes later. The post-anaesthesia validation study tested patients at the point of discharge from recovery [26]. A positive result still requires clinical assessment, because residual anaesthetic and sedative effects may contribute to the findings. In this case the drowsiness persisted two hours later, and

urgent assessment identified hypoxia. The anaesthetic history suggests a possible cause, but the score is scored as found. Follow the delirium tool specified by the local recovery-room pathway. The pre-operative o provides useful baseline information.

Case 10: “He’s just in pain”

Mr J, 82, day 1 after emergency laparotomy, is restless, pulling at his catheter and trying to climb out of bed. The working explanation at the desk is pain. He answers three AMT4 questions, one incorrectly, then loses engagement and gives no answer to the fourth. When asked to do Months Backwards, he understands the request but refuses to start. The night note records him settled until 2am, then “increasingly agitated”. A bladder scan shows retention around a blocked catheter; morphine doses have crept up.

Item 1: 4 (clearly agitated). **Item 2: 2** (one incorrect answer and one unanswered after engagement is lost). **Item 3: 1** (understands but refuses to start). **Item 4: 4. Total: 11/12.**

Comment. Pain and delirium are not alternative explanations. Undertreated pain, urinary retention and opioid exposure can each cause or worsen delirium. A positive 4AT does not prove that agitation is delirium, but this combination of acute change, abnormal alertness and cognitive failure demands clinical assessment. Here it prompted review of the blocked catheter, the analgesic plan and the drug chart.

Case 11: Day two, quietly worse

Mrs K, 79, is two days after bowel resection. She is not agitated; she is vague. She answers slowly, misses the year, and gives up months backwards at October, apologising. Her daughter, visiting, says she seems “flat and far away”, quite unlike yesterday.

Item 1: 0 (alert throughout, though slow). **Item 2: 1. Item 3: 1. Item 4: 4. Total: 6/12.**

Comment. Postoperative delirium peaks on days one to three, and much of it looks like this: nothing dramatic, just a positive score made

up of several small findings. Daily 4ATs for the first several post-operative days in older patients catch these presentations while they are still mild, when cause-hunting (here: opioid dose, low sodium, poor sleep) has the most to offer.

HIP FRACTURE AND ORTHOGERIATRICS

Case 12: Positive at presentation, and what it means

Mr L, 86, arrives with a fractured hip after a fall at home. He is alert. On the AMT4 he makes two errors; he attempts Months Backwards but stops at November. His son describes ten days of increasing muddle since a chest infection.

Item 1: 0. Item 2: 2. Item 3: 1. Item 4: 4. Total: 7/12.

Comment. In the UK national registry study of 18,040 hip fracture patients, around one in five had a positive 4AT at presentation; positive scores were associated with higher mortality, longer stays and a lower likelihood of returning home [17]. The score therefore marks a higher risk patient, rather than determining an individual prognosis. It should be handed to the whole team: the anaesthetic team when planning perioperative care, orthogeriatrics when identifying and treating causes and preventing worsening, the surgical team, and the family when discussing risk and likely course. A positive 4AT should inform care but should not, by itself, delay hip-fracture surgery. It is also the baseline against which post-operative scores will be read. This is not a form to be completed before the operation. The result changes how the perioperative team manages him, so it is part of the operation being done properly.

Case 13: Pre-op zero, post-op six

Mrs M, 81, scored 4AT 0 before surgery for her hip fracture. On post-operative day 1 she is alert but scores one AMT4 error, stops months backwards at September, and the nurses report she was disorientated overnight, settling by morning.

Item 1: 0. Item 2: 1. Item 3: 1. Item 4: 4 (overnight change, still evident within 24 hours). **Total: 6/12.**

Comment. Incident delirium was suspected within a day and subsequently confirmed on clinical assessment; the pre-operative o made the acute change unmistakable. Anaemia and constipation were identified and treated as likely contributors. The practical standard on this pathway is a 4AT at presentation and daily for the first days after surgery; her case shows the payoff. Item scores plus a one-line reason make the record readable at tomorrow's review.

Case 14: Severe dementia, score of 4, probably no delirium

Mrs N, 96, lives in a nursing home with severe dementia. She is admitted after a fall; X-ray shows a pubic ramus fracture. She is calm and awake, says her first name, and responds inconsistently to simple requests; AMT4 and months backwards cannot be meaningfully attempted. The home's senior carer, phoned, is definite: "this is exactly her usual self, she walked with her frame yesterday and ate breakfast this morning."

Item 1: 0. Item 2: 2 (untestable). **Item 3: 2** (untestable). **Item 4: 0** (informant reports no change). **Total: 4/12.**

Comment. A score of 4 indicates possible delirium, but the collateral history makes severe dementia without current delirium the likelier explanation here. This is the known specificity trade-off discussed in Chapter 4: untestable ratings protect drowsy delirious patients at the cost of flagging some people with severe dementia at baseline. The score still prompts clinical assessment of the fracture, pain and other possible contributors. It also forced the baseline call to the care home and puts her on the radar for monitoring, since severe dementia and a fresh fracture confer high delirium risk. So the score was recorded, the reasoning behind it was written down, and she was kept under review.

INTENSIVE CARE AND HIGH DEPENDENCY

Case 15: The right tool for the tube

Mr O, 68, is ventilated in ICU after emergency orthopaedic surgery. The surgical team asks for “a 4AT for baseline”. He cannot speak; assessment while ventilated belongs to the ICU’s own instruments. The unit’s CAM-ICU is positive, and clinical assessment confirms delirium. Three days later, sedative infusions have stopped, he is extubated, fully awake and able to sustain conversation, and is transferring from ICU to the surgical ward. On the AMT4 he gives his age and date of birth correctly but misses the place and year. He recites December, November and October, then stops. The ward team confirms that this cognitive change arose during the current admission and remains present.

Item 1: 0. Item 2: 2. Item 3: 1. Item 4: 4. Total: 7/12.

Comment. The 4AT was not designed for intubated patients; the CAM-ICU and ICDSC were designed exactly for them. While the patient remains in critical care, use the unit’s validated protocol. Extubation alone is not the signal to change instruments. The 4AT is most useful here once the patient can engage and at transition to ward care, where it gives the receiving team a familiar item-by-item description of persisting delirium. The transition assessment helps prevent delirium disappearing into the handover gap.

Case 16: High dependency, not ventilated

Mrs P, 71, is in a high dependency unit (HDU) with sepsis, on high-flow oxygen, not sedated. She is drowsy, rousable to voice, drifting mid-sentence. Cognitive testing cannot be started. Her husband says that four days ago she was fully alert, conversing normally and managing as usual.

Item 1: 4. Item 2: 2. Item 3: 2. Item 4: 4. Total: 12/12.

Comment. A four-centre Korean study found that the 4AT had a sensitivity of 74% and specificity of 95% against a DSM-5 reference in critically ill patients [38]. This supports its possible use in non-ventilated critical illness but does not supersede national or local guid-

ance. NICE recommends CAM-ICU or ICDSC rather than the 4AT in critical care [12]; units should use their agreed critical-care pathway. Whichever instrument is used, “she’s too sick to assess” should not be the end of it.

PALLIATIVE CARE AND HOSPICE

Case 17: Frightened, and findable

Mr Q, 66, has metastatic prostate cancer and is admitted to the hospice for symptom control. He repeatedly scans the room, startles markedly at ordinary sounds and cannot settle; he asks the nurse to make the insects on the wall go away. He answers the AMT4 with one error and stops Months Backwards at October. His wife says the fear started two days ago, around the time his opioid dose rose.

Item 1: 4 (hypervigilant). **Item 2: 1.** **Item 3: 1.** **Item 4: 4.** **Total: 10/12.**

Comment. A hospice validation study found that the 4AT performed strongly in hospice inpatients (sensitivity 89%, specificity 94%) [39]. The positive score prompted clinical assessment and a proportionate search for causes, including medication toxicity, dehydration, infection, constipation, urinary retention and metabolic disturbance. In this case the recent opioid increase was reviewed, opioid rotation was considered and calcium was checked. The hallucinations were explained as a recognised symptom of delirium, often related to illness or medication, rather than evidence that he was “going mad”, and his fear was treated actively. Some delirium in palliative care is reversible, particularly when related to drugs or dehydration; reversibility becomes less likely as death approaches.

Case 18: The last days

Mrs R, 84, is dying of heart failure; the agreed focus is comfort. Over 36 hours she has become deeply drowsy, briefly rousable to voice, occasionally plucking at the sheet, and no longer able to speak with her family. A brief 4AT is completed from observation and history, without burdening her with

repeated cognitive prompts: markedly reduced alertness (4), untestable on both cognitive items (2 + 2), and acute change (4). Total 12.

Comment. Whether the assessment was worth doing depends on whether it changes anything about her care. Here a single brief assessment helps name possible delirium and guides a proportionate, comfort-focused review. Check for contributors whose treatment might improve comfort, such as pain, urinary retention, constipation or medication effects. The score does not oblige anyone to investigate, catheterise, transfer or repeatedly retest. Explain the change to the family and treat distress. Assessment should serve the agreed goals of care.

Case 19: Recovery, and the conversation afterwards

Mr S, 72, developed delirium in the hospice after a chest infection. Serial 4ATs, undertaken while assessment remained tolerable and useful, showed: Day 1, clearly hypervigilant (Item 1: 4), one AMT4 error (Item 2: 1), at least seven Months Backwards correct (Item 3: 0), and clear acute change (Item 4: 4), total 9; Day 2, Item 1: 0, Item 2: 1, Item 3: 1, Item 4: 4, total 6; Day 4, Item 1: 0, Item 2: 1, Item 3: 0, Item 4: 4, total 5; and Day 5, Item 1: 0, Item 2: 1, Item 3: 0, Item 4: 0, total 1, with the single AMT4 error confirmed as his usual baseline and no acute change or fluctuation evident during the previous 24 hours. When clear, he is quiet, then asks the nurse whether he had “gone mad for a few days”. They talk. He remembers the ward as a ship’s engine room and his son as a stranger in a doorway.

Comment. Repeated 4ATs can help assess recovery in a hospice patient when testing is tolerated and the result informs care. A short, honest conversation afterwards, with written information for him and his son, can be highly valuable. It is worth asking a patient who has recovered what they remember of it, and then explaining to them what actually happened.

CARE HOMES

Case 20: Three weeks of fading, two days of drowsiness

Mr T, 92, lives in a residential home with Alzheimer's dementia; he normally walks with assistance, eats well, and enjoys company. Staff describe three weeks of increasing fatigue and, in the last 48 hours, real drowsiness, poor drinking, and reduced response. The visiting doctor finds him asleep; he wakes to voice but repeatedly drifts off during answers and needs prompting to re-engage. AMT4: knows the place and his age; cannot give year or date of birth. Months Backwards: "December", nothing further.

Item 1: 4. Item 2: 2. Item 3: 1. Item 4: 4. Total: 11/12.

Comment. A sudden deterioration on top of a slow decline is still a sudden deterioration, and Item 4 asks about the last two weeks; the last 48 hours clearly qualify. In care homes, staff who know the resident and have access to current baseline records can provide particularly strong collateral history; do not assume this automatically where staff are unfamiliar with the resident or records are sparse. The decision that followed, hospital or home-based treatment, was made with his family. What the score did was make the decision an urgent one, and make sure it was taken with the right information and with the family involved.

Case 21: Sundowning still needs explaining

Mrs U, 87, moderate dementia, has begun shouting and pacing every evening for four days, and her nights have inverted. At the 11am assessment she is calm and alert. AMT4: two errors, in line with her usual state per staff. Months Backwards: declines to try. Staff are clear the evening behaviour is new.

Item 1: 0 (at the time of assessment). Item 2: 2. Item 3: 1 (refusal). Item 4: 4 (new, fluctuating). Total: 7/12.

Comment. "She sundowns" describes a pattern; it does not explain a new change. A new evening-worse disturbance arising over

four days meets the Item 4 rule, although the positive 4AT still requires clinical assessment because fluctuation can also occur in dementia, particularly dementia with Lewy bodies. Constipation and a recently started bladder anticholinergic were identified as likely contributors. The case also shows the Item 3 boundary: a patient who verbally refuses to attempt Months Backwards scores 1; score 2 only when the patient cannot start because they are unwell, drowsy or inattentive.

Case 22: Home from hospital but not back to baseline

Mr V, 90, returns to his care home after a four-week admission with pneumonia complicated by delirium. The delirium began during the first week of the admission. During the final two weeks in hospital and the week since his return, staff have observed no further fluctuation or significant new mental-state change, although he remains less sharp than before admission. The home's nurse performs a 4AT: alert (0), two AMT4 errors (2), Months Backwards to September (1), and no change or fluctuation arising within the preceding two weeks and still evident during the last 24 hours (0). Total 3.

Comment. This is the sort of situation where a score in the 1 to 3 range is genuinely useful. This may be his new baseline; delirium can also take weeks to resolve; and dementia that was not previously recognised may be becoming apparent. What is needed here is a plan rather than a diagnosis: document the trajectory, watch for renewed fluctuation, repeat the 4AT if anything changes, and consider cognitive review after the acute episode has settled, often at six to twelve weeks, if this would change care and is consistent with the person's goals. What must be avoided is a dementia diagnosis being recorded while he is still recovering from delirium.

The timing matters here, so it is worth setting out how Item 4 is scored. Item 4 is governed by its precise wording: score 4 when a significant change or fluctuation arose within the preceding two weeks and remains evident in the last 24 hours. Here the original change began more than two weeks ago and there has been no signif-

icant new change or fluctuation during that window, so Item 4 scores 0. Persisting or resolving delirium nevertheless remains clinically possible despite a total of 3; the 4AT is not a rule-out test for prolonged delirium. If evidence is unavailable or uncertain, record that limitation, continue seeking collateral information and reassess rather than automatically converting uncertainty into 4 points.

COMMUNITY, PRIMARY CARE AND AMBULANCE

Case 23: Found on the floor, nobody to ask

A paramedic crew attends Mr W, 85, found on his kitchen floor by a neighbour. He is drowsy, rousing to voice and mumbling; cognitive testing cannot be started. No family is reachable. The neighbour has met him only briefly and cannot give a reliable mental-state baseline. The crew's system shows no care package and a repeat prescription collected in person last week, but no contemporaneous description of his cognition.

Item 1: 4. Item 2: 2. Item 3: 2. Item 4: 4 (no reliable evidence yet about the timing of mental-state change but it is a reasonable clinical judgement that Mr W is not normally like this). **Total: 12/12.**

Comment. This is a genuine no-informant situation. Items 1 to 3 already produce a strongly positive score, so the urgent response does not depend on guessing Item 4. However, in the present case the evidence strongly suggests that Mr W is different from his normal self. It is good practice to continue seeking collateral information from family, primary-care records and anyone who knows him, but 4AT Item 4 can be scored positive in this kind of situation.

Case 24: The kitchen table assessment

Her GP visits Mrs X, 88, at home; her daughter phoned the surgery saying her mother has been "away with the fairies" for ten days, worse in the evenings. At 3pm Mrs X is charming and plausible. AMT4: one error (the year). Months Backwards: confident start, reaches August, then stalls and

laughs it off. The daughter's account is specific. In the last ten days her mother has mistaken the hall cupboard for the bathroom, left a pile of letters unposted, and boiled a saucepan dry; all three are new, and the bathroom error happened again the previous night.

Item 1: 0. Item 2: 1. Item 3: 1. Item 4: 4. Total: 6/12.

Comment. Social fluency can survive during a lucid interval in delirium; the daughter's specific account of change, including an event in the last 24 hours, is more informative than the superficially reassuring conversation in front of the doctor. A positive 4AT at home requires prompt clinical assessment and a clear pathway. Do not diagnose a urinary infection from confusion or a urine dipstick alone. In this case she also has new dysuria and frequency, a urine sample is sent, and she is clinically dehydrated. Her observations are stable, there is no focal neurological deficit or other immediate red flag, she can drink and take oral treatment, and her daughter can stay with her. Home treatment is therefore agreed, with explicit deterioration advice and review the following day. Admission would be appropriate if the cause, physiological stability, oral intake, supervision or ability to review were less secure. Cognitive follow-up should be considered after recovery if concerns persist or the daughter describes decline predating these ten days.

NEUROLOGY, SENSORY IMPAIRMENT AND DIAGNOSTIC LIMITS

Case 25: Aphasia is not abnormal alertness

Mr Z, 78, is on day 2 after a left middle cerebral artery infarct causing severe aphasia. On admission he was awake, settled and able to follow some demonstrated one-step commands and indicate basic needs using gesture and a communication board. Overnight he became intermittently drowsy, pulled off his monitoring leads and then appeared clearer again. This morning he is awake and not agitated. A speech and language therapist helps establish the best available communication method, but he cannot give reliable answers to the AMT4 or demonstrate the Months Backwards

sequence. His family and the nurses agree that the overnight drowsiness and behavioural change were new and distinct from his stable language deficit.

Item 1: 0 (alertness normal during the present assessment). **Item 2: 2** (untestable despite communication support). **Item 3: 2** (untestable despite a supported attempt). **Item 4: 4** (new fluctuation beyond the established stroke-related aphasia). **Total: 8/12.**

Comment. Aphasia must not be scored as abnormal alertness: an awake patient who cannot understand or produce language still scores 0 on Item 1. Equally, the cognitive items should not be abandoned without trying reasonable communication support and, where possible, involving speech and language therapy. The aphasia is not the evidence for delirium here. The new drowsiness and behavioural fluctuation on top of it are. The positive score is not a diagnosis. In acute stroke, deterioration also demands urgent assessment for stroke extension, haemorrhagic transformation, seizure, hypoxia, hypoglycaemia and other neurological or medical complications. The 4AT must not delay the stroke pathway. Document what the patient could communicate, how testing was attempted and which findings were genuinely new.

Case 26: Check the hearing aids before scoring

Mrs AA, 91, has severe bilateral hearing impairment and is admitted after a fall. From the end of the bed she appears to ignore questions, but she watches the speaker closely and is not drowsy. Before starting the 4AT, the nurse checks both hearing aids, replaces a flat battery, reduces the background noise and sits facing her. With the aids working, Mrs AA engages readily. She makes one AMT4 error, giving last year as the current year, and recites December, November and October before stopping. Her daughter describes new night-time disorientation during the previous two days, including mistaking the bedroom for the bathroom the previous night.

Item 1: 0 (awake and attentive; deafness is not abnormal alertness). **Item 2: 1** (one error). **Item 3: 1** (started but completed fewer than

seven months). **Item 4: 4** (specific acute change, evident during the last 24 hours). **Total: 6/12.**

Comment. The unsuccessful exchange from the end of the bed was not a valid cognitive assessment. Reasonable adjustment is part of doing the 4AT properly: check that hearing aids are present, switched on and working; reduce competing noise; face the patient; speak clearly; and, if vision and literacy permit, write the questions down. Do not score deafness as drowsiness or automatically declare the cognitive items untestable. After adjustment, this patient still has a positive 4AT and needs clinical assessment of possible delirium and its causes, including the circumstances and consequences of the fall. Record the adjustment used so that later scores are comparable.

Case 27: Poor sight and longstanding visual experiences

Mrs AB, 88, has severe macular degeneration and is admitted with cellulitis. She tells the admitting nurse that she sometimes sees patterned figures in the room. Her daughter confirms that these experiences have occurred intermittently for many months, are unchanged, and that Mrs AB normally recognises that they are not real. The assessment is conducted in good light, with her usual glasses on; the clinician identifies herself and describes the unfamiliar surroundings. Mrs AB is fully alert, answers all four AMT4 questions correctly and recites the months backwards from December to June. Her daughter reports no significant new change or fluctuation in alertness, thinking or behaviour during the previous two weeks or the last 24 hours.

Item 1: 0. Item 2: 0. Item 3: 0. Item 4: 0. Total: 0/12.

Comment. The 4AT contains no drawing, reading or other inherently visual task, so severe visual impairment does not by itself make the patient untestable. These longstanding, unchanged visual experiences associated with sight loss are compatible with Charles Bonnet syndrome and do not satisfy Item 4's acute-change rule. However, neither the baseline history nor a score of 0 gives permanent reassurance. New or markedly changed hallucinations, loss of insight,

altered alertness, inattention or functional deterioration should prompt reassessment and a search for delirium and other causes. Do not dismiss new symptoms as poor eyesight. Equally, a longstanding visual phenomenon on its own is not evidence of delirium.

Case 28: Parkinson's disease – slow is not drowsy

Mr AC, 78, has Parkinson's disease. At baseline he moves slowly and speaks quietly but is fully orientated; his wife says he does not usually hallucinate. Over 36 hours he has intermittently seen children in the room and repeatedly lost the thread of conversations, with lucid intervals. At assessment he has masked facies, hypophonia and marked bradykinesia, but is fully awake, follows the conversation and is not agitated. He answers all four AMT4 questions correctly. Months Backwards: December, November, October, then he stalls. His wife confirms that this is a definite change, with another episode overnight.

Item 1: 0. Item 2: 0. Item 3: 1. Item 4: 4. Total: 5/12.

Comment. Parkinsonian slowness, reduced facial expression and quiet speech do not by themselves constitute abnormal alertness; Mr AC remains fully alert, so Item 1 scores 0. The positive total comes from a clearly dated, fluctuating change in attention and perception. New hallucinations in Parkinson's disease may reflect delirium, medication effects or Parkinson's-related psychosis; the 4AT does not decide between them. Clinical assessment should look for acute illness, pain, dehydration, constipation, urinary retention and recent medication changes. Antiparkinsonian medicines must be given at their individually prescribed times and must not be stopped abruptly. Do not reach reflexively for haloperidol or another dopamine-blocking drug; if medication is being considered for severe distress or immediate risk, follow Parkinson's-specific guidance and seek specialist advice.

Case 29: Lewy body dementia – change from a fluctuating baseline

Mrs AD, 82, has established dementia with Lewy bodies. A new care worker requests an assessment after Mrs AD describes small animals in the sitting room and alternates between lively conversation and muddled answers. Her daughter and care plan confirm that the same non-distressing visual hallucinations and fluctuations have occurred several times a week for 18 months. There has been no new drowsiness, loss of function, reduced intake or worsening of cognition or behaviour during the preceding two weeks, including the previous 24 hours. At assessment she is fully alert. She makes two AMT4 errors, consistent with her documented baseline, and gives December, November and October before stopping.

Item 1: 0. Item 2: 2. Item 3: 1. Item 4: 0. Total: 3/12.

Comment. The word ‘fluctuation’ in Item 4 does not mean that every fluctuation scores 4. The finding must represent a significant acute change or fluctuation arising during the preceding two weeks and still evident within the last 24 hours. Mrs AD’s hallucinations and variable cognition are longstanding features of her dementia, not a new change. Her total of 3 flags cognitive impairment but does not suggest current delirium. Dementia with Lewy bodies does not protect against delirium: a new or clearly worse pattern should trigger clinical assessment, renewed collateral history and reassessment. Where it remains difficult to distinguish delirium from baseline fluctuation in dementia with Lewy bodies, senior clinical input may be needed. Familiar, non-distressing hallucinations should not automatically be treated with medication. Antipsychotics can worsen motor features and, in some people with Lewy body dementia, cause severe sensitivity reactions.

Case 30: A score of 0 does not close the question

Mrs AE, 83, is admitted with fever, cough and poor fluid intake. A triage note says that she seemed ‘a little vague earlier’, but gives no examples; her family cannot initially be reached. At 11am she is fully alert, answers the

AMT4 correctly and recites Months Backwards correctly to June. There is not yet reliable evidence of a significant acute mental-state change. Because the collateral history is thin and the earlier concern remains unexplained, the clinician documents the limitation and plans reassessment rather than treating the score as the end of the enquiry.

Initial assessment: Item 1: 0. Item 2: 0. Item 3: 0. Item 4: 0.
Total: 0/12.

Her daughter later contacts the ward. Mrs AE is normally independent and has never behaved in this way before, but during the previous 48 hours she twice believed that her house was a hotel, including an episode during the preceding night. A nurse observes another brief episode of disorientation that afternoon. At 6pm Mrs AE is again lucid: alert, AMT4 correct and Months Backwards correct to June. The new collateral and observed fluctuation now establish Item 4.

Repeat assessment: Item 1: 0. Item 2: 0. Item 3: 0. Item 4: 4.
Total: 4/12.

Comment. Both scores were correctly calculated from the evidence available at the time. Acute illness and vague concern do not themselves justify four points for Item 4; the later specific collateral history does. A score of 0 makes delirium less likely but does not definitively exclude it, particularly during a lucid interval or when information about recent change is incomplete. Reassessment should be prompted by continuing clinical concern or new evidence, not performed mechanically. Immediate clinical assessment and treatment of physiological instability, hypoglycaemia, focal neurological signs or other emergencies must never wait for the score.

INDEX OF LESSONS

Unstable patients still getting a score: Cases 1, 5, 16, 23, 25

Language and interpreters: 2

Fluctuation and repeat testing: Cases 3, 8, 21, 28, 30

Agitation and hyperactive presentations: Cases 4, 10

Hypoactive delirium: Cases 5, 11, 18

Delirium superimposed on dementia: Cases 6, 20, 21
Severe dementia and possible false positives: 14
Recovery tracking: Cases 7, 19, 22
Perioperative timing and baselines: Cases 9, 12, 13
No or incomplete collateral history: Cases 23, 30
Prognosis and pathways (hip fracture): Cases 12, 13
ICU boundaries and transitions: Cases 15, 16
Palliative and end-of-life use: Cases 17, 18, 19
Borderline scores and cognitive follow-up: Cases 7, 22, 24
Talking to patients and families: Cases 17, 19, 24
Stroke and aphasia: 25
Hearing impairment and reasonable adjustment: 26
Visual impairment and longstanding hallucinations: 27
Parkinson's disease: 28
Dementia with Lewy bodies: Cases 21, 29
A zero despite continuing suspicion: Cases 3, 30

IMPLEMENTING THE 4AT, AND KNOWING WHETHER IT IS WORKING

Choosing a good tool is a small part of the work, and much the easiest part. Instruments with strong validation data under research conditions can nonetheless work poorly once they reach routine clinical practice: completed patchily, scored wrongly, modified without revalidation, or abandoned within the year. This gap between research and routine performance has been documented across delirium tools generally, the 4AT included [41,51]. The barriers are real and sometimes formidable, but they are knowable, and most have workflow solutions rather than motivational ones. This chapter is about implementation, drawn from studies, audits, quality improvement projects, and from seeing many services succeed and fail at this over two decades.

WHAT IMPLEMENTATION INVOLVES

A defined trigger and a defined response. Decide exactly when the 4AT happens (for example: all patients 65 and over within 24 hours of admission; daily for days 1 to 5 after hip fracture surgery; on any positive SQiD) and exactly what a positive score triggers (who assesses, within what time, against which pathway). If nothing happens as a

result of a positive score, you end up with a lot of recorded scores and staff who stop believing the exercise is worth anything.

Where the 4AT lives in the record. Paper forms can disappear; the 4AT works best embedded in the admission documentation and the electronic record, with the item wording on the form, automatic totalling, and the score visible where decisions get made (handover sheets, ward lists, ED tracking screens). The 4AT is free to embed in electronic health record (EHR) and electronic medical record (EMR) systems, with no licensing or permission required; thousands of hospitals and systems across the world have done so.

Education: start with why it matters. The 4AT needs no formal training to administer, but a short information session can enhance the quality of its use as well as helping staff understand what they are doing and why. Start by explaining what delirium is and why it matters (start with patient stories; staff who grasp how distressing delirium is need little further persuading), the item rules, the untestable logic, and what happens after a positive. Thirty to sixty minutes is enough. Build it into induction, because staff rotate, and have ward champions refresh it. Videos and case sets exist for exactly this (Chapter 15). The website www.deliriumacademy.com, written by the author, offers free training on delirium.

Named ownership. Somewhere, a person or small group owns delirium detection: reviewing the numbers, answering questions from the wards, and keeping the subject on the agenda once the launch enthusiasm has worn off. Improvements that depend entirely on one champion melt away when the champion moves on; ownership needs to be institutional, with leadership that keeps asking for the results.

TWO KEY REAL-WORLD IMPLEMENTATION NUMBERS

Once running, a detection process is checked with two simple figures, both extractable from any electronic record.

Completion rate. What proportion of the intended patients actu-

ally had a 4AT? Published large-scale experience shows what is achievable. Strikingly, completion reached 91% across 18,040 hip fracture patients in national registry data [17], with near-90% rates in Norwegian ward and stroke-unit programmes [33,44]. Rates far below these usually mean workflow problems (the form is hard to find, the trigger is ambiguous, nobody looks at the result) rather than staff indifference, and they are usually solved by changing the workflow. A New Zealand study of barriers found exactly this pattern: staff supported the idea while having difficulties with the practicalities [51].

Positive score rate. Of the 4ATs done, what proportion scored 4 or more, and does that match the delirium rates in your population? Around 15 to 25% of older acute admissions should score positive; the national hip fracture data ran at 21% [17], and a systematic review of detection tools used at scale found that 4AT implementations reported positive rates consistent with expected delirium prevalence, and that some 4AT programmes achieved among the highest completion rates in the review [41]. Completion rate is largely a property of the programme rather than of the instrument. The 91% figure above comes from a pathway in which the field is a mandated national audit item, so it should be read as evidence that high completion is achievable with the 4AT, not that the 4AT produces it. If a service is assessing acute elderly admissions and only 3% score positive, the explanation is almost never that the hospital has no delirium in it. It usually means the well patients are the ones being tested, or the scoring has drifted, or the form is being ticked without the assessment being done. Comparing the expected positive rate with the observed one is cheap to do and usually tells you more than anything else you could measure.

Track both numbers over time and feed them back to the wards by name. Teams that are finding a lot of delirium are usually the ones doing the assessment properly, and it is worth saying so. A rise in the detection rate is something to be pleased about.

COMMON CAUSES OF FAILURE, AND THEIR FIXES

These kinds of failures are seen repeatedly across services. The form completed but the score ignored: the fix is to the pathway and to where the score is displayed, rather than to the staff. Untestable misused as unable-to-be-bothered, scoring drowsy patients or by skipping items: fix with the one-line rule that untestable feeds the score, taught through cases 1 and 5. The tool quietly modified (an item dropped, the threshold moved) to make workflow easier: modified tools lose their evidence base (Chapter 12); resist, or use limited-period daily testing as designed. Assessment only of the confused-looking: systematic testing exists precisely because a lot of delirium is not visible from the nursing station. And launch-and-forget: the programmes that endure make the two numbers a standing item at governance meetings and re-teach at every staff intake.

A NOTE ON SCALE

None of this requires a research team. One person with access to the record system can pull the completion and positive rates each month. One interested member of staff per ward can keep the standard up, and an hour of teaching at each intake keeps the knowledge fresh. The published experience, from single wards [33] to entire national pathways [17], shows the same shape: clarity about when and what next, embedding in the record, brief education with stories, and visible feedback of the two numbers. Services that get those four things right do find delirium. Choosing a tool and stopping there does not achieve much on its own.

THE EVIDENCE BASE

This chapter gives clinical readers the evidence in usable form: what kind of studies exist, the headline accuracy figures, the real-world performance data, and the limitations. The full study list, updated as new work appears, can be found here: www.the4AT.com/4atevidence.

THE SHAPE OF THE EVIDENCE

The 4AT has two evidence bases, and the distinction is important, because few tools have a similar level of evidence of both types.

The first is diagnostic test accuracy: studies in which patients received the 4AT and, separately and blind, a reference-standard delirium assessment against formal criteria. There are 33 such studies involving over 6,000 patients, spanning emergency departments, acute medical and geriatric wards, surgical and post-anaesthesia settings, cardiac surgery, stroke units, hospices, skilled nursing facilities, intensive care and traumatic brain injury, in some 20 languages. Considered as an individual instrument, the 4AT has one of the largest formal validation bases among delirium tools and, in the 2021

review, roughly double the number of patients studied for the next most-studied individual instrument [23].

The second is implementation at scale: studies of what happens when the 4AT is used in routine care by ordinary staff, measured in completion rates, positive score rates and predictive relationships with outcomes, in datasets from the tens of thousands upward. This second evidence base matters because of what has happened to previous tools. Instruments with strong validation data under research conditions can nonetheless work poorly once they reach routine clinical practice: completed patchily, scored wrongly, modified without revalidation, or abandoned within the year. This gap between research and routine performance has been documented across delirium tools generally, the 4AT included [41,51]. Accuracy studies on their own only tell you that a tool can work under study conditions. You need the implementation data as well before you know whether it works on an ordinary ward.

DIAGNOSTIC ACCURACY: THE HEADLINE FIGURES

The 2021 systematic review and meta-analysis led by Zoë Tieges pooled 17 studies (3,702 observations): sensitivity 88% (95% confidence interval (CI) 80 to 93) and specificity 88% (95% CI 82 to 92) [23]. The review used a bivariate random-effects model and reports summary sensitivity and specificity rather than a pooled area under the curve. That review was conducted by members of the group that developed the 4AT, including the present author. Developer-led evaluations tend on average to report more favourable results than independent ones, so the independent replication that follows carries particular weight.

An independent meta-analysis by Hou and colleagues, conducted without involvement from the 4AT developers (11 studies, 2,789 participants), reached almost identical pooled figures: sensitivity 87%, specificity 87% [52].

Subsequent validation studies have been broadly consistent with

these estimates, though not uniformly so. Two of the larger recent studies reported lower sensitivities than the pooled figure: 76% in the UK multicentre randomised comparison [11] and 74% in non-ventilated critically ill patients in South Korea [38]. Both reported high specificity.

A separate diagnostic test accuracy study in geriatric medical inpatients evaluated the 4AT alongside other delirium assessment tools [53]. The 4AT was most feasible and balanced sensitivity (86.7%) with reasonable specificity (69.5%).

For a two-minute test requiring no training, those are strong numbers, and their consistency across settings and languages is the more important finding. Selected individual results, all against formal reference standards:

Bellelli 2014 [9]: *Acute geriatric and rehabilitation wards, Italy.* Sensitivity 89.7%, specificity 84.1%; AUC 0.93; AUC 0.89 in patients with dementia

Shenkin 2019 [11]: *ED and acute medicine, UK, 785 patients analysed.* Sensitivity 76%, specificity 94%; AUC 0.90; randomised comparative diagnostic accuracy study with CAM (see Chapter 11)

De 2017 [24]: *Geriatric wards, Australia; 80% probable dementia; interpreters used.* Sensitivity 87%, specificity 80%; in probable dementia 86% and 71%

Saller 2019 [26]: *Post-anaesthesia care unit, Germany.* Validated for recovery-room use after surgery

Arnold 2024 [39]: *Two hospices, Scotland, 134 analysed.* Sensitivity 89%, specificity 94%; AUC 0.97

Sim 2024 [38]: *Four ICUs, South Korea, non-ventilated critically ill.* Sensitivity 74%, specificity 95%; AUC 0.88

The De study is worth noting separately. In a population where four in five patients had probable dementia and many needed interpreters, the 4AT retained its sensitivity, with lower specificity in the dementia subgroup [24].

REAL-WORLD PERFORMANCE

Five large studies anchor this half of the evidence.

260,200 older-adult admissions, eight hospitals in the USA (Pan 2025). Katy Agarwal and George Taffet introduced the 4AT to their hospital system in 2015, using Days Backwards in place of Months Backwards. The 4AT was used twice daily; delirium was associated with mortality, longer stays and institutional discharge [54].

18,040 hip fracture patients, UK national registry (Penfold 2025). 4AT completion 91% in routine care; 21% positive; positive scores independently associated with mortality, longer stay and new residential care [17].

82,770 emergency admissions, two Scottish centres (Anand 2022). Routine admission 4AT scores linked to outcomes: positive scores associated with higher mortality, longer stays and less time at home [16].

75,221 admissions (Penfold 2025, Age Ageing). 4AT scores at admission, including the 1 to 3 band, were strongly associated with an existing or subsequent dementia diagnosis. This is the evidence behind the tool's second job as a cognitive impairment flag [19].

Systematic review of tools at scale (Penfold 2024). Across studies of detection tools used in large routine populations, the 4AT showed high completion rates and positive score rates matching expected delirium prevalence; several comparators showed positive rates far below plausible prevalence, implying missed cases at scale [41].

SMALLER IMPLEMENTATION STUDIES fill in the picture: nurse-led 4AT use on acute geriatric wards and stroke units in Norway with completion around 90% [33,44], ED use in suspected sepsis [32], a digital delirium pathway in an English trust improving detection and outcomes [55], and a UK-wide survey finding the 4AT the most used tool in acute hospitals [10]. Barriers research documents where completion falters and why [51].

Recovery assessment. McCartney and colleagues (2025) validated

repeated 4ATs for assessing recovery from delirium in older inpatients, formalising a use many services had already adopted [27].

LIMITATIONS

No delirium tool is diagnostic, including this one; the diagnosis is always clinical, and the 4AT's output is possible delirium. Specificity falls slightly in severe dementia, for the structural reason explained in Chapter 4; users must know this, and the manual says so wherever it is relevant. Sensitivity is not 100%: mild delirium presenting in a lucid interval can score below 4, so persisting suspicion should lead to a repeat assessment rather than being taken as reassurance. The evidence in ventilated ICU patients is absent by design (Chapter 6), the evidence in some settings (community, paramedic use) rests on extrapolation and pragmatic experience more than formal validation, and most studies come from high-income health systems. Item 2 and Item 3 performance is affected by education and language, handled by sensible administration (Chapter 3) but not eliminated. Finally, the 4AT is an assessment tool. It will not improve outcomes unless someone acts on the result.

THE 4AT COMPARED WITH OTHER TOOLS

More than 60 delirium assessment instruments exist. This chapter compares the 4AT with the ones users actually ask about, on the dimensions that decide real-world performance: time, training, testability of sick patients, what the output means, and evidence in practice.

CONFUSION ASSESSMENT METHOD (CAM)

The Confusion Assessment Method (1990) [56] operationalises delirium into four features rated present or absent, and the tool's manual strictly requires a preceding cognitive test and structured interview, and formal training in rating the features. Done as designed, by trained assessors, it has performed well in many diagnostic test accuracy studies.

The most direct comparison available is a multicentre study in which patients were randomised to receive either the 4AT or the CAM, each followed by a blinded reference-standard assessment (Shenkin 2019; 785 patients analysed, emergency department and acute medicine, DSM-IV reference standard). In that study the 4AT had a sensitivity of 76% (95% CI 61 to 87) and a specificity of 94% (95%

CI 92 to 97); the CAM had a sensitivity of 40% (95% CI 26 to 57) and a specificity of 100% (95% CI 98 to 100) [11].

Two points of context are needed. First, the CAM was administered as its manual specifies, by research staff trained from the CAM instruction manual, following a structured interview and a set of cognitive tests. This was not a shortened or unofficial CAM, and the result cannot be explained away as misuse. Second, the CAM's sensitivity here is markedly lower than its pooled sensitivity of 82% (95% CI 69 to 91) across 22 studies in an earlier meta-analysis, in which pooled specificity was 99% [57].

What the study supports is a claim about detection strategies rather than about instruments in the abstract: in an unselected acute medical and emergency population, a pathway built on the 4AT identified substantially more delirium than a pathway built on the CAM, with only a small reduction in specificity. For a condition whose principal failure mode is non-detection, that is the trade most services should want.

In routine practice the CAM is not always used as its manual specifies. Local electronic-record implementations have been described in which the four features are recorded as a symptom checklist without the preceding cognitive testing and structured interview; this is an unofficial process: it is not what the CAM's authors describe, and its accuracy has not been established [50]. Where detection tools have been studied at scale in routine care, some services show high completion alongside positive-score rates well below plausible delirium prevalence, which implies that cases are being missed while the record looks complete [41]. Recording a negative delirium screen in a patient who has delirium is clinically serious, because it converts an omission into positive false reassurance; at scale it does so systematically. The point is not that the CAM performs badly when used as designed. It does not. A tool whose validity depends on preceding steps is vulnerable when those steps are dropped, and every service using such a tool should audit specifically for this.

In 2023 the UK's NICE guideline replaced the short CAM with the

4AT as the recommended tool outside critical care, while retaining the CAM-ICU and ICDSC in critical care. NICE judged the 4AT “among the most accurate of the tools reviewed, quick and simple to use”, with “a broader range of evidence to support it” [12].

THE CAM FAMILY

There are numerous variants of the CAM. Three commonly used outside the ICU are the Brief Confusion Assessment Method (bCAM) [58], the 3-Minute Diagnostic Interview for CAM-defined Delirium (3D-CAM) [59], and the Ultra-Brief Confusion Assessment Method (UB-CAM) [60]. All are well validated in research studies and perform similarly to the 4AT. The 3D-CAM has also been assessed in a systematic review and meta-analysis [61], and the UB-CAM has been evaluated in routine implementation [62]. They each require training to use. They output a binary delirium yes/no result rather than a graded score with a cognitive-impairment band, as the 4AT does. Currently there is less information about clinical implementation at scale than there is with the 4AT. A service currently using one of these tools successfully, with proven completion and positive-score rates [41], has no urgent reason to switch. The lack of need for training for the 4AT may be a consideration, however. A service choosing a new tool may find the 4AT the quicker and simpler option to deploy. The results of forthcoming studies, along with a better understanding of real-world implementation, will be informative.

OBSERVATIONAL MONITORING SCALES: NURSING DELIRIUM SCREENING SCALE (NU-DESC), DELIRIUM OBSERVATION SCREENING SCALE (DOSS)

The Nu-DESC [48] and DOSS [49] are shift-based nursing observation tools. For each, a set of delirium features is scored from behaviour observed during routine care over the shift; if the relevant points threshold is crossed, delirium is considered possible. As

Chapter 7 argued, they answer a different question (has something changed, and does it warrant assessment?) and are best used alongside the 4AT rather than instead of it. Their sensitivity as stand-alone detectors is variable, and neither should be the final word before a delirium diagnosis is applied or withheld. One disadvantage of tools scored at the end of a shift is that delirium present earlier may be detected late.

RECOGNIZING ACUTE DELIRIUM AS PART OF YOUR ROUTINE (RADAR)

RADAR is a three-item observational screen completed by nursing staff during medication rounds [63]. It records drowsiness, difficulty following instructions and slowed movements. Any positive item makes the screen positive. It requires no direct questioning or baseline information and took about seven seconds in the original study. In some clinical workflows, RADAR is used for ongoing monitoring, with a positive result triggering a 4AT for a more definitive assessment.

THE SINGLE QUESTION IN DELIRIUM (SQiD)

The Single Question in Delirium (SQiD) is asked of a relative, friend or carer who knows the patient. Its original wording was: “Do you think [patient’s name] has been more confused lately?” [22]. Minor variations are used in practice, including “Do you feel that [patient’s name] has been more confused lately?” and versions that ask whether the person is “more confused or drowsy than usual”. This effectively is incorporated into the 4AT as part of Item 4. As a stand-alone measure it is a monitoring signal, not an assessment: a positive SQiD should trigger a 4AT. In the UK NHS, a commonly used inpatient monitoring sequence combines NEWS2 [47] with the SQiD and uses the 4AT after a positive trigger.

INTENSIVE CARE UNIT (ICU) TOOLS

The CAM-ICU [36] is built for ventilated and non-verbal patients, using picture-recognition and letter tasks that can be answered by nodding or squeezing a hand. The Intensive Care Delirium Screening Checklist (ICDSC) [37] works differently: it is a checklist completed by nursing staff from all the information available over a shift, including the patient's speech and behaviour where these can be observed, rather than a single bedside test. Both remain the standard inside intensive care, with the 4AT taking over when speech returns and at step-down (Chapter 6). The Korean multicentre study suggests the 4AT is a reasonable alternative in non-ventilated critically ill patients [38].

COGNITIVE TESTS ARE NOT DELIRIUM TESTS

A recurring source of muddle: the Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), 10-item Abbreviated Mental Test (AMT10) and similar instruments measure cognition, not delirium. A delirious patient will usually score poorly on them, but so will a person with dementia; none of them assesses acute change, fluctuation or arousal, and none should be used to rule delirium in or out. Their proper role is later: baseline cognition once the acute episode has resolved (Chapter 5). The 4AT cognitive items may offer an option for Mentation within the Age-Friendly Health Systems 4Ms Framework [4,19]. Further information: <https://www.the4at.com/4at-cognitive-assessment-dementia>. Cognitive testing alone is not a delirium assessment.

WHICH TOOL TO CHOOSE

For episodic assessment in general settings, including front door, perioperative care, care homes, hospices and the community, the 4AT is a reasonable default: quick to deploy, with no formal training barrier, a score for every patient including those who cannot engage,

a graded output, and an exceptionally broad validation base together with the largest published implementation datasets, now reflected in national guidelines [12,15]. For continuous inpatient monitoring, use a light observational process (for example, SQiD, NEWS2-linked, NuDESC, DOSS) with the 4AT as the follow-on assessment. Inside ventilated intensive care, use the ICU-specific tools. Whatever the choice, check the two numbers from Chapter 9; a tool that is not being completed, or that never scores positive, is not working, however good its published accuracy.

VERSIONS, TRANSLATIONS AND TERMS OF USE

VERSIONS

The 4AT was first published at www.the4AT.com in 2011 after iterative clinical testing. Version 1.2 dates from 2014. Validation evidence applies to its stable wording; altered versions may perform differently. Future changes will appear on the website, which hosts the definitive form.

MODIFIED 4AT

Some services replace Months Backwards with Days Backwards. Most validation studies used Months Backwards, but the Tieges 2021 review included Texas data from the modified 4AT [23]. Pan and colleagues documented its use at very large scale [54]. A patient unable to undertake either task receives the same score of 2, which can contribute to a positive 4AT, so the effect on delirium scoring is likely to be minimal. Days Backwards is easier, however, so interpretation of total 4AT scores from 1 to 3 may differ. Full evidence, history, and updates: <https://www.the4at.com/modified-4at>.

TRANSLATIONS

The 4AT exists in more than 20 languages, including Arabic, Chinese, Czech, Danish, Dutch, Finnish, French, German, Hebrew, Icelandic, Italian, Japanese, Korean, Norwegian, Polish, Portuguese, Russian, Spanish, Swedish, Thai and Turkish, among others.

Translations follow forward and back translation with clinical review. The Thai version was developed and validated in hospitalised older people with acute medical illnesses [64].

Other translations with published validation studies include Italian [9], French [29], Spanish [30], German (post-operative) [26], Dutch [65], Swedish [66], Japanese [67], Korean, including a remote-use adaptation studied during COVID-19 [68], and Traditional Chinese [46] (see Chapter 10; the full current list is at www.the4AT.com/4at-translations). Available translations are free to download from the website. If a language is missing, contact the team via the website; new translations are welcomed and supported.

TERMS OF USE

The 4AT is free to use, copy and distribute for clinical, educational and research purposes. It will always be free. No permission, registration or payment is required. This includes reproduction in local guidelines and teaching materials, and building the 4AT into electronic health record systems, commercial or otherwise, without licensing fees. Content is licensed under CC BY 4.0; the source (www.the4AT.com) should be attributed.

When referencing the tool in publications, cite the original validation study [9] and the website www.the4AT.com, plus the meta-analysis [23] where relevant.

RELATED TOOLS: THE SQID AND THE 4-DSD

The Single Question in Delirium (SQiD) asks someone who knows the patient whether they think the person has been more confused lately [22]. In practice a drowsiness variant is often used: “is she more confused or drowsy than usual?” This is not the original wording but suits bedside use. It is the smallest useful delirium instrument, and its logic is folded into 4AT Item 4. As a stand-alone it works as a monitoring question, asked once or twice a shift on at-risk wards, or put to a family member at any encounter; a positive answer is a signal to do a 4AT, not a diagnosis. It costs nothing to use, and it asks the person who actually knows the patient’s baseline.



THE 4-DSD ADDRESSES the hardest corner of the field: detecting delirium superimposed on dementia in people with moderate to severe dementia, which is where the specificity limitation described in Chapter 4 applies. It is a 4-item tool developed by Alessandro Morandi, with colleagues including the 4AT authors, drawing on observational assessment of arousal and attention rather than testing

that severe dementia makes impossible; the validation study is published [69] and the tool and instructions are free at www.the4AT.com/4-dsd. For services with substantial severe-dementia populations, care homes above all, the 4-DSD is the natural companion to the 4AT.

Also part of the family in a looser sense: months backwards [6,7] and the AMT4 [4,5] exist as stand-alone bedside tests with their own literatures, and understanding them repays the 4AT user; both are discussed in Chapter 3.

FREQUENTLY ASKED QUESTIONS

These are the questions users actually send, lightly organised. Where an answer needs a chapter, the chapter is pointed to rather than repeated.

GENERAL

What is the 4AT? A two-minute bedside assessment tool for delirium and cognitive impairment, with four items: Alertness, AMT4, Attention (months backwards), and Acute change or fluctuating course. Scores of 4 or more suggest possible delirium. It requires no special training and is free to use.

What does the name mean? The 4 items each begin with A; AT stands for Assessment Test.

Who developed it, and when? Alasdair MacLullich (University of Edinburgh), with Tracy Ryan and Helen Cash (NHS Lothian), first published in 2011; the current version 1.2 dates from 2014. Updates are the responsibility of the author. The development story is Chapter 1.

Who is it for? Adults (18 and over) in any clinical setting outside the ICU, with the strongest evidence and use in people 65 and over. It is not validated in children and adolescents.

Who can administer it? Any healthcare professional: nurses, doctors of any specialty, therapists, pharmacists, paramedics, physician associates, and students within supervision norms. Local policy governs unregistered staff. The constraint is not that the tool is hard to use; it is making sure the result reaches someone who will act on it.

Is it really free? In our EHR too? Yes, and yes. No permission, registration or licence fee, including for integration into electronic record systems (Chapter 12). There will never be a fee for using the 4AT.

Which guidelines recommend it? In the UK, NICE (2023 update; the 4AT replaced the short CAM as the recommended tool outside critical care) and SIGN (2019). Elsewhere, examples include the European Society of Anaesthesiology and Intensive Care postoperative guideline (2024 update), the Australian Delirium Clinical Care Standard, and American Psychiatric Association guidance, along with many regional and hospital pathways [12,15,70–72]. More information on the 4AT's presence in guidelines can be found here: <https://www.the4at.com/4at-in-guidelines>.

PRACTICAL USE

How long does it take? Under two minutes in routine practice.

Do I need training? No formal training is required; that is a design feature, and validation studies have generally used assessors given brief orientation to the tool rather than a training course. A short orientation (this manual's Chapters 3 and 4, the website videos, the cases) improves consistency and is recommended for services (Chapter 9).

When should it be done? At presentation to acute care in older adults; before and after surgery; on transitions of care; whenever delirium is suspected, including on family concern; after a positive monitoring signal (SQiD, NEWS2-linked observation); and daily to track recovery during an episode (Chapters 5 to 7).

How often can it be repeated? Daily, or somewhat more often around clinical change, for limited periods: perfectly appropriate,

and validated for recovery tracking [27]. Multiple times a day for weeks is not appropriate: patient and staff burden and practice effects on the cognitive items degrade the result. That job belongs to monitoring instruments (Chapter 7).

The patient was asleep. Score the sleep? Attempt to wake with speech or a gentle touch. Normal sleep from which the patient wakes to normal alertness scores 0 on Item 1. A patient who cannot be roused to sustained normal alertness is showing altered arousal: score accordingly (Chapter 3).

There is no informant. Can I still complete Item 4? Often, yes: notes, letters, prior assessments and the clinical course count as evidence, and an obvious deterioration from a documented independent baseline can be scored 4 with clinical judgement (see Case 23). Where no evidence of change exists either way, score 0 and remember the score's meaning is provisional on that gap; the other items still provide routes to a positive (Chapters 3 and 4).

The patient speaks no English. Use an interpreter; ask the AMT4 in translation; do months backwards in the patient's language; or use one of 20+ official translations. Language alone never makes a patient untestable (Chapter 3, Case 2).

Can it be done remotely? Telephone and video assessment have been used in practice and studied, including a Korean remote adaptation during COVID-19 [68]. Items 2 and 3 transmit well; Item 1 needs a good video view; Item 4 is unaffected. Reasonable where in-person assessment is impossible, with the mode documented.

Should I do a 4AT on someone who obviously has delirium? Yes, briefly: the score documents severity-relevant features, time-stamps the episode, and gives later assessments a comparator. It only takes two minutes.

SCORING

Why does one abnormal item (1 or 4) score enough to flag delirium? Because both findings, clearly altered alertness and acute change or fluctuation, are individually highly specific for delirium in

general settings. The 4AT's overall accuracy with this weighting has been demonstrated repeatedly (Chapter 10). The weighting itself was a clinical design decision rather than a statistically derived one: validation studies test the tool as a whole, not the contribution of individual items.

When is a patient “untestable”? When their clinical state (drowsiness, severe inattention, inability to produce speech) prevents a meaningful attempt at the AMT4 or months backwards. Refusing or abandoning months backwards while engaged scores 1, not 2. Untestable ratings feed the score by design: no patient is Unable to Assess (Chapters 3 and 4).

Does a score of 4 or more diagnose delirium? No. It means possible delirium; the diagnosis is clinical, made after assessment. Equally, a score below 4 with strong clinical suspicion means you should reassess rather than treat the score as reassurance (Chapters 4 and 5).

What exactly does 1 to 3 mean? Cognitive errors without current evidence of acuity: possible cognitive impairment. It calls for baseline informant history, delirium vigilance, and cognitive follow-up after the acute episode (Chapter 5, Cases 22 and 24). It is not a dementia diagnosis.

The patient has severe dementia and scored 4 from two untestable items. Delirium? It may be either, and the score cannot tell you which. What it does do is make you go and find out. Establish the baseline from someone who knows the person, look for change, and monitor. The 4-DSD is designed for this population (Chapters 4 and 13, Case 14).

Can the 4AT measure severity? It was not designed as a severity scale, though higher scores tend to accompany more severe presentations, and repeated scores validly track recovery [27]. For formal severity measurement in research, dedicated instruments exist, including the Delirium Rating Scale-Revised-98 (DRS-R98) [73], Memorial Delirium Assessment Scale (MDAS) [74], and Confusion Assessment Method-Severity (CAM-S) [75].

Does the 4AT diagnose dementia? No, but its scores flag the

possibility usefully: in 75,221 admissions, higher admission 4AT scores were strongly associated with dementia diagnoses [19]. Formal cognitive assessment happens after recovery, not during the acute episode (Chapter 5).

EVIDENCE

How well validated is it? 33 diagnostic accuracy studies, more than 6,000 patients, pooled sensitivity and specificity both around 88% [23,52]. Considered as an individual instrument, the 4AT has one of the largest formal validation bases among delirium tools (Chapter 10).

And in the real world? The largest implementation datasets published for any delirium tool: 91% completion and 21% positive rates in 18,040 hip fracture patients, with positive scores associated with outcomes [17]; 82,770 admissions linking scores to mortality and time at home [16]; systematic review evidence that its routine-practice positive rates match true prevalence where other tools' do not [41] (Chapters 9 and 10).

Why do some electronic-record delirium screens detect so little delirium? The commonest explanation, where it has been investigated, is that the tool is being completed without the assessment steps its accuracy depends on, most often a preceding cognitive test and structured interview that the manual requires but that a screen embedded in the record makes easy to skip. Check your positive-score rate against expected prevalence (Chapter 9); if it is far too low, look first at how the tool is actually being administered rather than at its published accuracy. See Chapter 11.

SITUATIONS

Intensive care? Ventilated patients: use CAM-ICU or ICDSC. Non-ventilated critically ill: 4AT performs acceptably [38]. Step-down and transitions: 4AT (Chapter 6, Cases 15 and 16).

Parkinson's disease and Lewy body dementia? Fluctuation and

hallucinations occur in these diseases without delirium; scores need senior clinical interpretation, informed by the person's own pattern [45] (Chapter 6).

Dying patients? The 4AT works in hospices with strong accuracy [39]; assessment is worth doing if the result would change something, such as comfort measures, what you tell the family, or a decision about a reversible cause. If it would not, do not do it (Chapter 6, Cases 17 to 19).

Documentation: what should a good 4AT entry look like? Four item scores with one-line reasons, the total, and the action: "4AT: 1=4 (drifting off repeatedly), 2=2 (untestable), 3=2 (untestable), 4=4 (daughter: clear change over 2 days). Total 12. Treating as delirium; causes review started." (Chapter 4.)

A colleague disagrees with my score. Good: the disagreement is usually about the underlying judgement (was alertness clearly abnormal? does the history amount to acute change?), which is clinical and worth having. Score what you observed, record the reasoning, and let the clinical assessment resolve it. Inter-rater reliability of the 4AT has been reported as good where it has been studied.

Where do I ask something this list missed? www.the4AT.com, which also hosts the current form, translations, videos, cases and the updated FAQ.

RESOURCES

The website. www.the4AT.com hosts the definitive form, translations, the user guide, FAQ, evidence lists and updates. Everything is free without registration.

Videos. A three-minute walkthrough of the 4AT and a ten-minute teaching video with three scored case examples are linked from the user guide page; both suit induction teaching.

Cases. Chapter 8 of this manual; the worked examples at www.the4AT.com/4atcases; and multiple further interactive scored cases linked from that page.

Delirium care. www.the4AT.com/deliriumguide provides a concise guide; the SIGN delirium guideline (2019) [15]; the NICE delirium guideline (2023 update) [12].

Related tools. The 4-DSD for delirium in moderate to severe dementia, with instructions, at www.the4AT.com/4-dsd [69].

Delirium education: www.deliriumacademy.com (by the author) provides free educational modules on delirium and links to other resources. www.deliriumday.com also provides useful resources.

Delirium assessment tools: www.deliriumtools.com (by the author);

NIDUS: deliriumnetwork.org.

Information for families: www.deliriumsupport.com (by the author) can help point families towards useful support once delirium is diagnosed.

Going deeper. For the science of delirium: the Nature Reviews Disease Primers review [13]. *Delirium*: the field's focused journal, created and backed by the European Delirium Association, American Delirium Society, and Australasian Delirium Association, and World Delirium Awareness Day each March (www.deliriumday.com).

ABOUT THE AUTHOR

Alasdair MacLulich is Professor of Geriatric Medicine at the University of Edinburgh and a consultant physician at the Royal Infirmary of Edinburgh, working in acute geriatric medicine and orthogeriatrics. He qualified in medicine at Edinburgh in 1993, with an intercalated BSc in Psychology, and completed a PhD on cognitive ageing, a lectureship, and a Medical Research Council (MRC) Clinician Scientist Fellowship before appointment as Professor in 2008. His research centres on delirium: clinical assessment, neuropsychology and pathophysiology, with related work in dementia, frailty and hip fracture care. He co-founded the European Delirium Association in 2006 and has served as its President, co-founded the Scottish Delirium Association in 2011 and World Delirium Awareness Day in 2016, co-chaired the SIGN delirium guideline (2019), chairs the Scottish Hip Fracture Audit, and is Editor in Chief of the journal *Delirium* (www.deliriumjournal.com). He led the design of the 4AT in 2011 and remains responsible for it. See www.alasdairmaculich.com for more details.

ACKNOWLEDGEMENTS

The author thanks Tracy Ryan and Helen Cash, both then of NHS Lothian, who collaborated in the design and early testing of the 4AT and are co-authors of the original validation study [9]. He also thanks the colleagues who generously contributed time and expertise during the initial development and testing of the 4AT: Atul Anand,

Daniel Davis, Fergus Doubal, Duncan Forsyth, Kyle Gibson, Aditya Hosakere, Bree Johnston, Colin MacDonald, Katie Marwick, Johanna Paddison, Emma Reynish, Jeyakumar Selwyn and Susan Shenkin. Thanks also to the researchers worldwide who have validated and studied the 4AT independently, to the clinical teams who have built it into everyday care, and to the patients and families whose experiences of delirium are the reason the tool exists.

REFERENCES

1. Tieges Z, McGrath A, Hall RJ, MacLullich AMJ. Abnormal level of arousal as a predictor of delirium and inattention: an exploratory study. *Am J Geriatr Psychiatry*. 2013;21(12):1244–1253. doi:10.1016/j.jagp.2013.05.003.
2. Hall R, Stiobhairt A, Allerhand M, MacLullich AMJ, Tieges Z. The Observational Scale of Level of Arousal: a brief tool for assessing and monitoring level of arousal in patients with delirium outside the ICU. *Int J Geriatr Psychiatry*. 2020;35(9):1021–1027. doi:10.1002/gps.5324.
3. European Delirium Association; American Delirium Society. The DSM-5 criteria, level of arousal and delirium diagnosis: inclusiveness is safer. *BMC Med*. 2014;12:141. doi:10.1186/s12916-014-0141-2.
4. Schofield I, Stott DJ, Tolson D, McFadyen A, Monaghan J, Nelson D. Screening for cognitive impairment in older people attending accident and emergency using the 4-item Abbreviated Mental Test. *Eur J Emerg Med*. 2010;17(6):340–342. doi:10.1097/MEJ.0b013e32833777ab.
5. Swain DG, Nightingale PG. Evaluation of a shortened version of the Abbreviated Mental Test in a series of elderly patients. *Clin Rehabil*. 1997;11(3):243–248. doi:10.1177/026921559701100308.
6. Meagher J, Leonard M, Donoghue L, O'Regan N, Timmons S, Exton C, et al. Months backward test: a review of its use in clinical studies. *World J Psychiatry*. 2015;5(3):305–314. doi:10.5498/wjp.v5.i3.305.
7. O'Regan NA, Ryan DJ, Boland E, Connolly W, McGlade C, Leonard M, et al. Attention! A good bedside test for delirium? *J Neurol Neurosurg Psychiatry*. 2014;85(10):1122–1131. doi:10.1136/jnnp-2013-307053.
8. Tieges Z, Stiobhairt A, Scott K, Suchorab K, Weir A, Parks S, et al. Development of a smartphone application for the objective detection of attentional deficits in delirium. *Int Psychogeriatr*. 2015;27(8):1251–1262. doi:10.1017/S1041610215000186.
9. Bellelli G, Morandi A, Davis DHJ, Mazzola P, Turco R, Gentile S, et al. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing*. 2014;43(4):496–502. doi:10.1093/ageing/afu021. Corrigendum: *Age Ageing*. 2015;44(1):175. doi:10.1093/ageing/afu181.
10. Tieges Z, Lowrey J, MacLullich AMJ. What delirium detection tools are used in routine clinical practice in the United Kingdom? Survey results

- from 91% of acute healthcare organisations. *Eur Geriatr Med*. 2021;12(6):1293–1298. doi:10.1007/s41999-021-00507-2.
11. Shenkin SD, Fox C, Godfrey M, Siddiqi N, Goodacre S, Young J, et al. Delirium detection in older acute medical inpatients: a multicentre prospective comparative diagnostic test accuracy study of the 4AT and the Confusion Assessment Method. *BMC Med*. 2019;17(1):138. doi:10.1186/s12916-019-1367-9.
 12. National Institute for Health and Care Excellence. Delirium: prevention, diagnosis and management in hospital and long-term care. London: NICE; 2010 [updated 2023]. (Clinical guideline CG103).
 13. Wilson JE, Mart MF, Cunningham C, Shehabi Y, Girard TD, MacLulich AMJ, et al. Delirium. *Nat Rev Dis Primers*. 2020;6(1):90. doi:10.1038/s41572-020-00223-4.
 14. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5-TR. 5th ed, text rev. Washington (DC): American Psychiatric Association Publishing; 2022. doi:10.1176/appi.books.9780890425787.
 15. Scottish Intercollegiate Guidelines Network. Risk reduction and management of delirium. Edinburgh: SIGN; 2019. (SIGN publication no. 157).
 16. Anand A, Cheng M, Ibitoye T, MacLulich AMJ, Vardy ERLC. Positive scores on the 4AT delirium assessment tool at hospital admission are linked to mortality, length of stay and home time: two-centre study of 82,770 emergency admissions. *Age Ageing*. 2022;51(3):afac051. doi:10.1093/ageing/afac051.
 17. Penfold RS, Farrow L, Hall AJ, Clement ND, Ward K, Donaldson L, et al. Delirium on presentation with a hip fracture is associated with adverse outcomes: a multicentre observational study of 18,040 patients using national clinical registry data. *Bone Joint J*. 2025;107-B(4):470–478. doi:10.1302/0301-620X.107B4.BJJ-2024-1164.R1.
 18. Hshieh TT, et al. Effectiveness of multicomponent nonpharmacological delirium interventions: a meta-analysis. *JAMA Intern Med*. 2015;175(4):512–520. doi:10.1001/jamainternmed.2014.7779.
 19. Penfold RS, Bowman E, Vardy ERLC, Sampson EL, Anand A, Guthrie B, et al. Using scores from the 4AT delirium detection tool as an indicator of possible dementia: a study of 75 221 older adult hospital admissions. *Age Ageing*. 2025;54(6):afafi44. doi:10.1093/ageing/afafi44.
 20. Tiegies Z, Quinn T, MacKenzie L, Davis D, Muniz-Terrera G, MacLulich AMJ, et al. Association between components of the delirium syndrome

- and outcomes in hospitalised adults: a systematic review and meta-analysis. *BMC Geriatr.* 2021;21(1):162. doi:10.1186/s12877-021-02095-z.
21. Hasemann W, Duncan N, Clarke C, Nouzova E, Süßenbach LM, Keerie C, et al. Comparing performance on the Months of the Year Backwards test in hospitalised patients with delirium, dementia, and no cognitive impairment: an exploratory study. *Eur Geriatr Med.* 2021;12(6):1257–1265. doi:10.1007/s41999-021-00521-4.
 22. Sands MB, Dantoc BP, Hartshorn A, Ryan CJ, Lujic S. 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(6):561–565. doi:10.1177/0269216310371556.
 23. Tiegies Z, MacLulich AMJ, Anand A, Brookes C, Cassarino M, O'Connor M, et al. Diagnostic accuracy of the 4AT for delirium detection in older adults: systematic review and meta-analysis. *Age Ageing.* 2021;50(3):733–743. doi:10.1093/ageing/afaa224.
 24. De J, Wand APF, Smerdely PI, Hunt GE. Validating the 4A's test in screening for delirium in a culturally diverse geriatric inpatient population. *Int J Geriatr Psychiatry.* 2017;32(12):1322–1329. doi:10.1002/gps.4615.
 25. Richardson SJ, Davis DHJ, Bellelli G, Hasemann W, Meagher D, Kreisel SH, et al. Detecting delirium superimposed on dementia: diagnostic accuracy of a simple combined arousal and attention testing procedure. *Int Psychogeriatr.* 2017;29(10):1585–1593. doi:10.1017/S1041610217000916.
 26. Saller T, MacLulich AMJ, Schäfer ST, Crispin A, Neitzert R, Schüle C, et al. Screening for delirium after surgery: validation of the 4 A's test (4AT) in the post-anaesthesia care unit. *Anaesthesia.* 2019;74(10):1260–1266. doi:10.1111/anae.14682.
 27. McCartney H, Noble E, Thompson K, Fomina K, Guevara LM, Davis DHJ, et al. Validation of the 4AT for assessing recovery from delirium in older hospital patients. *Age Ageing.* 2025;54(6):afaf166. doi:10.1093/ageing/afaf166.
 28. O'Sullivan D, Brady N, Manning E, O'Shea E, O'Grady S, O'Regan N, et al. Validation of the 6-Item Cognitive Impairment Test and the 4AT test for combined delirium and dementia screening in older emergency department attendees. *Age Ageing.* 2018;47(1):61–68. doi:10.1093/ageing/afx149.
 29. Gagné AJ, Voyer P, Boucher V, Nadeau A, Carmichael PH, Pelletier M, et al. Performance of the French version of the 4AT for screening the elderly for delirium in the emergency department. *CJEM.* 2018;20(6):903–910. doi:10.1017/cem.2018.367.

30. Morales-Puerto M, Ruiz-Díaz M, García-Mayor S, León-Campos Á, Morales-Asencio JM, Canca-Sánchez JC, et al. Spanish transcultural adaptation of the 4AT score for the evaluation of delirium in the emergency department: a prospective diagnostic test accuracy study. *BMC Nurs.* 2024;23(1):101. doi:10.1186/s12912-023-01638-6.
31. Soler-Sanchis A, Martínez-Arnau FM, Sánchez-Frutos J, Pérez-Ros P. The 4AT scale for rapid detection of delirium in emergency department triage. *Front Med (Lausanne).* 2024;11:1345983. doi:10.3389/fmed.2024.1345983.
32. Myrstad M, Kuwelker K, Haakonsen S, Valebjørg T, Langeland N, Kittang BR, et al. Delirium screening with 4AT in patients aged 65 years and older admitted to the emergency department with suspected sepsis: a prospective cohort study. *Eur Geriatr Med.* 2022;13(1):155–162. doi:10.1007/s41999-021-00558-5.
33. Myrstad M, Watne LO, Johnsen NT, Børs-Lind E, Neerland BE. Delirium screening in an acute geriatric ward by nurses using 4AT: results from a quality improvement project. *Eur Geriatr Med.* 2019;10(4):667–671. doi:10.1007/s41999-019-00215-y.
34. Chang Y, Ragheb SM, Oravec N, Kent D, Nugent K, Cornick A, et al. Diagnostic accuracy of the “4 A’s Test” delirium screening tool for the postoperative cardiac surgery ward. *J Thorac Cardiovasc Surg.* 2023;165(3):1151–1160.e8. doi:10.1016/j.jtcvs.2021.05.031.
35. Penfold RS, Hall AJ, Anand A, Clement ND, Duckworth AD, MacLulich AMJ. Delirium in hip fracture patients admitted from home during the COVID-19 pandemic is associated with higher mortality, longer total length of stay, need for post-acute inpatient rehabilitation, and readmission to acute services. *Bone Jt Open.* 2023;4(6):447–456. doi:10.1302/2633-1462.46.BJO-2023-0045.R1.
36. Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. Delirium in mechanically ventilated patients: validity and reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *JAMA.* 2001;286(21):2703–2710. doi:10.1001/jama.286.21.2703.
37. Bergeron N, Dubois MJ, Dumont M, Dial S, Skrobik Y. Intensive Care Delirium Screening Checklist: evaluation of a new screening tool. *Intensive Care Med.* 2001;27(5):859–864. doi:10.1007/s001340100909.
38. Sim JK, Chung K, Chung CR, Lee J, Hwang SY, Lee YS. Usefulness of the 4A’s test for detecting delirium in critically ill patients: a multicenter prospective observation study. *Intern Emerg Med.* 2024;19(7):1839–1846. doi:10.1007/s11739-024-03670-z.

39. Arnold E, Finucane AM, Taylor S, Spiller JA, O'Rourke S, Spenceley J, et al. The 4AT, a rapid delirium detection tool for use in hospice inpatient units: findings from a validation study. *Palliat Med.* 2024;38(5):535–545. doi:10.1177/02692163241242648.
40. Baird L, Spiller JA. A quality improvement approach to cognitive assessment on hospice admission: could we use the 4AT or Short CAM? *BMJ Open Qual.* 2017;6(2):e000153. doi:10.1136/bmjopen-2017-000153.
41. Penfold RS, Squires C, Angus A, Shenkin SD, Ibitoye T, Tiegies Z, et al. Delirium detection tools show varying completion rates and positive score rates when used at scale in routine practice in general hospital settings: a systematic review. *J Am Geriatr Soc.* 2024;72(5):1508–1524. doi:10.1111/jgs.18751.
42. Lees R, Corbet S, Johnston C, Moffitt E, Shaw G, Quinn TJ. Test accuracy of short screening tests for diagnosis of delirium or cognitive impairment in an acute stroke unit setting. *Stroke.* 2013;44(11):3078–3083. doi:10.1161/STROKEAHA.113.001724.
43. Infante MT, Pardini M, Balestrino M, Finocchi C, Malfatto L, Bellelli G, et al. Delirium in the acute phase after stroke: comparison between methods of detection. *Neurol Sci.* 2017;38(6):1101–1104. doi:10.1007/s10072-017-2832-x.
44. Ihle-Hansen H, Johnsen N, Jankowiak T, Hagberg G, Walle-Hansen MM, Landgraff I, et al. Delirium screening in a stroke unit by nurses using 4AT: results from a quality improvement project. *Nurs Open.* 2023;10(11):7431–7436. doi:10.1002/nop2.1856.
45. Lawson RA, Richardson SJ, Gerakios F, Yarnall AJ, Bate G, Wright L, et al. Detecting delirium in Parkinson's disease: an evaluation of diagnostic accuracy of bedside tools. *Age Ageing.* 2025;54(7):afaf197. doi:10.1093/ageing/afaf197.
46. Lai YH, Lin CJ, Su IC, Huang SW, Hsiao CC, Jao YL, et al. Clinical utility and performance of the Traditional Chinese version of the 4-As Test for delirium due to traumatic brain injury. *J Acad Consult Liaison Psychiatry.* 2025;66(2):130–138. doi:10.1016/j.jaclp.2024.12.005.
47. Royal College of Physicians. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. Updated report of a working party. London: Royal College of Physicians; 2017. ISBN 978-1-86016-682-2. Available from: <https://www.rcp.ac.uk/resources/national-early-warning-score-news-2/>
48. Gaudreau JD, Gagnon P, Harel F, Tremblay A, Roy MA. Fast, systematic, and continuous delirium assessment in hospitalized patients: the

- Nursing Delirium Screening Scale. *J Pain Symptom Manage*. 2005;29(4):368–375. doi:10.1016/j.jpainsymman.2004.07.009.
49. Schuurmans MJ, Shortridge-Baggett LM, Duursma SA. The Delirium Observation Screening Scale: a screening instrument for delirium. *Res Theory Nurs Pract*. 2003;17(1):31–50. doi:10.1891/rtnp.17.1.31.53169.
 50. MacLulich AMJ. The CAM-Lite: why this unofficial delirium screening tool falls dangerously short. *Delirium Words*. Published 26 April 2025; updated 19 July 2026. Accessed 23 July 2026. Available from: <https://deliriumwords.com/delirium-words-1/the-cam-lite-why-this-unofficial-delirium-screening-tool-falls-dangerously-short/>
 51. Alhaidari AAO, Matsis KP. Barriers to completing the 4AT for delirium and its clinical implementation in two hospitals: a mixed-methods study. *Eur Geriatr Med*. 2022;13(1):163–172. doi:10.1007/s41999-021-00582-5.
 52. Hou L, Zhang Q, Cao L, Chen M, Wang Q, Li Y, et al. Diagnostic accuracy of the 4AT for delirium: a systematic review and meta-analysis. *Asian J Psychiatr*. 2023;80:103374. doi:10.1016/j.ajp.2022.103374.
 53. Hendry K, Quinn TJ, Evans J, Scortichini V, Miller H, Burns J, et al. Evaluation of delirium screening tools in geriatric medical inpatients: a diagnostic test accuracy study. *Age Ageing*. 2016;45(6):832–837. doi:10.1093/ageing/afw130.
 54. Pan AP, Agarwal K, Potter T, Borei K, Nowakowski S, Bruce CR, et al. Social disparities, delirium occurrence, and related outcomes among hospitalized older adults. *J Am Geriatr Soc*. 2025;73(12):3729–3737. doi:10.1111/jgs.70137.
 55. Vardy E, Collins N, Grover U, Thompson R, Bagnall A, Clarke G, et al. Use of a digital delirium pathway and quality improvement to improve delirium detection in the emergency department and outcomes in an acute hospital. *Age Ageing*. 2020;49(4):672–678. doi:10.1093/ageing/afaa069.
 56. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med*. 1990;113(12):941–948. doi:10.7326/0003-4819-113-12-941.
 57. Shi Q, Warren L, Saposnik G, MacDermid JC. Confusion assessment method: a systematic review and meta-analysis of diagnostic accuracy. *Neuropsychiatr Dis Treat*. 2013;9:1359–1370. doi:10.2147/NDT.S49520.
 58. Han JH, Wilson A, Vasilevskis EE, Shintani A, Schnelle JF, Dittus RS, et al. Diagnosing delirium in older emergency department patients: validity and reliability of the Delirium Triage Screen and the Brief Confusion

- Assessment Method. *Ann Emerg Med.* 2013;62(5):457–465.
doi:10.1016/j.annemergmed.2013.05.003.
59. Marcantonio ER, Ngo LH, O'Connor M, Jones RN, Crane PK, Metzger ED, 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(8):554–561. doi:10.7326/M14-0865.
 60. Motyl CM, Ngo LH, Zhou W, Jung Y, Leslie DL, Boltz M, et al. Comparative accuracy and efficiency of four delirium screening protocols. *J Am Geriatr Soc.* 2020;68(11):2572–2578. doi:10.1111/jgs.16711.
 61. Ma R, Zhao J, Li C, Qin Y, Yan J, Wang Y, et al. Diagnostic accuracy of the 3-minute diagnostic interview for confusion assessment method-defined delirium in delirium detection: a systematic review and meta-analysis. *Age Ageing.* 2023;52(5):afado74. doi:10.1093/ageing/afado74.
 62. Marcantonio ER, Fick DM, Jung Y, Inouye SK, Boltz M, Leslie DL, et al. Comparative implementation of a brief app-directed protocol for delirium identification by hospitalists, nurses, and nursing assistants: a cohort study. *Ann Intern Med.* 2022;175(1):65–73. doi:10.7326/M21-1687.
 63. Voyer P, Champoux N, Desrosiers J, Landreville P, McCusker J, Monette J, et al. Recognizing acute delirium as part of your routine [RADAR]: a validation study. *BMC Nurs.* 2015;14:19. doi:10.1186/s12912-015-0070-1.
 64. Kuladee S, Prachason T. Development and validation of the Thai version of the 4 'A's Test for delirium screening in hospitalized elderly patients with acute medical illnesses. *Neuropsychiatr Dis Treat.* 2016;12:437–443. doi:10.2147/NDT.S97228.
 65. Pouw MA, Calf AH, Georg RR, de Rooij SE, ter Maaten JC, van Munster BC. Diagnostic accuracy of the Dutch version of the 4AT for delirium detection in a mixed patient population and setting. *Aging Clin Exp Res.* 2023;35(8):1705–1710. doi:10.1007/s40520-023-02447-2.
 66. Johansson YA, Tsevis T, Nasic S, et al. Diagnostic accuracy and clinical applicability of the Swedish version of the 4AT assessment test for delirium detection, in a mixed patient population and setting. *BMC Geriatr.* 2021;21(1):568. doi:10.1186/s12877-021-02493-3.
 67. Hasegawa T, Seo T, Kubota Y, Sudo T, Yokota K, Miyazaki N, et al. Reliability and validity of the Japanese version of the 4A's Test for delirium screening in the elderly patient. *Asian J Psychiatry.* 2022;67:102918. doi:10.1016/j.ajp.2021.102918.
 68. Hur HJ, Jang YN, Park HY, Lee YS, Ro DH, Kang B, et al. A prospective study of remote delirium screening using the modified K-4AT for COVID-19 inpatients. *Front Psychiatry.* 2022;13:976228. doi:10.3389/fpsy.2022.976228.

69. Morandi A, Grossi E, Lucchi E, Zambon A, Faraci B, Severgnini J, et al. The 4-DSD: a new tool to assess delirium superimposed on moderate to severe dementia. *J Am Med Dir Assoc.* 2021;22(7):1535–1542.e3. doi:10.1016/j.jamda.2021.02.029.
70. Aldecoa C, Bettelli G, Bilotta F, Sanders RD, Aceto P, Audisio R, et al. Update of the European Society of Anaesthesiology and Intensive Care Medicine evidence-based and consensus-based guideline on postoperative delirium in adult patients. *Eur J Anaesthesiol.* 2024;41(2):81–108. doi:10.1097/EJA.0000000000001876.
71. Australian Commission on Safety and Quality in Health Care. Delirium clinical care standard. Sydney: ACSQHC; 2021.
72. American Psychiatric Association. The American Psychiatric Association practice guideline for the prevention and treatment of delirium. 2nd ed. Washington (DC): American Psychiatric Association Publishing; 2025. doi:10.1176/appi.books.9780890428023.
73. Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N. Validation of the Delirium Rating Scale-revised-98: comparison with the Delirium Rating Scale and the Cognitive Test for Delirium. *J Neuropsychiatry Clin Neurosci.* 2001;13(2):229–242. doi:10.1176/jnp.13.2.229.
74. Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K, Passik S. The Memorial Delirium Assessment Scale. *J Pain Symptom Manage.* 1997;13(3):128–137. doi:10.1016/S0885-3924(96)00316-8.
75. Inouye SK, Kosar CM, Tommet D, Schmitt EM, Puelle MR, Saczynski JS, Marcantonio ER, Jones RN. The CAM-S: development and validation of a new scoring system for delirium severity in 2 cohorts. *Ann Intern Med.* 2014;160(8):526–533. doi:10.7326/M13-1927.

Delirium is common, serious and often missed. The 4AT is a rapid bedside test for possible delirium and cognitive impairment. It takes about two minutes and requires no special training.

This manual explains how to administer and score each of the four items, interpret the result, and decide what should happen next. It addresses the situations that most often cause uncertainty, including drowsiness, agitation, dementia, inability to engage, language differences, and incomplete collateral history. Thirty worked cases show how the scoring rules apply in emergency care, hospital wards, perioperative and critical care, hospices, care homes, and community practice. Further chapters cover teaching, use in practice, audit, the evidence base, and comparison with other delirium tools.

For nurses, doctors, allied health professionals, paramedics, educators, students, and quality-improvement teams.

ABOUT THE AUTHOR

Alasdair MacLulich is Professor of Geriatric Medicine at the University of Edinburgh and a consultant physician in acute geriatric medicine and orthogeriatrics at the Royal Infirmary of Edinburgh. His research focuses on delirium. He led the design of the 4AT in 2011.