A systematic review of studies reporting on neuropsychological and functional domains used for assessment of recovery from delirium in acute hospital patients

A systematic review of studies reporting on neuropsychological and functional domains used for assessment of recovery from delirium in acute hospital patients

Haruno McCartney[1*], Erin Noble[1*], Alasdair MJ MacLullich[1], Daniel HJ Davis[2], Jonathan Evans[3], Susan D Shenkin[1,4], Graciela Muniz-Terrera[5,6], Daisy Sandeman[7], and Zoë Tieges[1,8]

[1]Edinburgh Delirium Research Group, Ageing and Health, Usher Institute, University of Edinburgh, Scotland, UK.
[2]MRC Unit for Lifelong Health and Ageing at University College London, London, UK.
[3]School of Health and Wellbeing, University of Glasgow, Glasgow, UK.
[4]Advanced Care Research Centre, Usher Institute, University of Edinburgh, Scotland, UK.
[5]Edinburgh Dementia Prevention and Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK.
[6]Department of Social Medicine, Ohio University, USA.
[7]School of Health in Social Science, University of Edinburgh, Scotland, UK.
[8]Department of Computing, School of Computing, Engineering and Built Environment, Glasgow Caledonian University, Scotland, UK.

*HM and EN should be considered joint first authors

Corresponding author:
Dr Zoë Tieges
Email: zoe.tieges@ed.ac.uk
Address: Room S1642, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh, EH16 4SA

Word count: 4498

Acknowledgements: Special thanks to the academic support librarian at the University of Edinburgh, Ruth Jenkins, who helped to refine the search strategies.

Sponsors: This review is part of a project funded by the Dunhill Medical Trust (grant reference: RPGF1902\147).

ABSTRACT

Objectives: Assessing for recovery in delirium is essential in guiding ongoing investigation and treatment. Yet, there is little scrutiny and no research or clinical consensus on how recovery should be measured. We reviewed studies which used tests of neuropsychological domains and functional ability to track recovery of delirium longitudinally in acute hospital settings.
Methods/Design: We systematically searched databases (MEDLINE, PsycInfo, CINAHL, Embase, ClinicalTrials.gov, Cochrane Central Register of Controlled Trials), from inception to October 14th, 2022. Inclusion criteria were: adult acute hospital patients (≥18 years) diagnosed with delirium by a validated tool; 1+ repeat assessment using an assessment tool measuring domains of delirium/functional recovery ≤7 days from baseline. Two reviewers independently screened articles, performed data extraction, and assessed risk of bias. A narrative data synthesis was completed.
Results: From 6,533 screened citations, we included 39 papers (reporting 32 studies), with 2,370 participants with delirium. Studies reported 21 tools with an average of four repeat assessments including baseline (range 2-10 assessments within ≤7 days), measuring 15 specific domains. General cognition, functional ability, arousal, attention and psychotic features were most commonly assessed for longitudinal change. Risk of bias was moderate to high for most studies.
Conclusions: There was no standard approach for tracking change in specific domains of delirium. The methodological heterogeneity of studies was too high to draw firm conclusions on the effectiveness of assessment tools to measure delirium recovery. This highlights the need for standardised methods for assessing recovery from delirium.

Keywords:
Delirium recovery, delirium assessment, 4AT, neuropsychological domains, repeat assessment, longitudinal studies, functional recovery.

Key points:

1. Assessing for recovery is an essential part of delirium care, to guide clinicians' ongoing investigation and treatment and to provide accurate information to patients and carers.
2. It is unclear what assessment tools are used to measure delirium recovery on a longitudinal basis, specifically which neuropsychological and functional domains should be assessed.
3. A total of 21 assessment tools measuring 15 different symptom domains of delirium were used in the included studies, however, there was no standard approach for tracking change in these domains over time.
4. The symptom domains most frequently assessed for longitudinal change in the included studies were general cognition, functional ability, arousal, attention and psychotic features.

INTRODUCTION

Delirium is a severe neuropsychiatric syndrome with acute onset and fluctuating course, characterised by impairments in attention, level of arousal, other domains of cognition (e.g., visuospatial ability, orientation etc.), psychosis and mood changes. It can be triggered by illness, infection, surgery or drugs[1] and is prevalent in acute hospital settings, affecting ~25% of older hospitalised in-patients[2]. In addition to impairment in neuropsychological domains, delirium can also negatively impact functional ability[3,4]. It is associated with adverse long-term outcomes such as higher mortality and increased dementia risk[1]; outcomes of patients with persistent delirium are especially poor[5].

It is essential to assess for recovery from delirium to guide health professionals' investigation and treatment of delirium, to give accurate information to patients and relatives, and thus provide optimal care to patients. However, there is uncertainty over how delirium recovery should be assessed and there is no standardised approach in clinical practice. Several assessment tools have been developed and validated for the use of delirium detection, such as the 4AT[6], the Confusion Assessment Method (CAM)[7], the Single Question in Delirium (SQiD)[8], etc., but it is unclear whether these tools are appropriate to assess delirium recovery.

It is also unclear how different neuropsychological domains are affected throughout delirium recovery trajectories. Several assessment tools of specific symptom domains exist, including tests of attention (e.g., Months of the Year Backwards[9], Digit Span[10], DelApp[11]) and arousal (e.g., Richmond Agitation Sedation Scale (RASS)[12], the Observational Scale of Level of Arousal (OSLA)[13]), although none were explicitly designed for repeated administration to assess patients with delirium. Understanding how neuropsychological domains of delirium fluctuate in the context of recovery will be helpful in determining which domains should be the focus of assessment, complementing other approaches such as psychiatric assessment. As well as specific neuropsychological domains, it is also of interest to evaluate more general tests of cognition such as the Mini-Mental State Examination (MMSE)[14] as these kinds of tests are commonly used to track cognitive functioning in practice.

To date, no systematic literature review has focused on the assessment of delirium recovery. A previous systematic review focused on definitions of delirium recovery, finding a wide range of approaches (various cut-off points, percentage reductions on delirium severity scales, or one or more days of 'negative' in dichotomous delirium present/absent scales)[15]. They concluded that consistent terminology in defining delirium recovery was required, and that cognitive recovery should be central to defining delirium.

We aimed to identify what longitudinal assessment tools are used within a clinically applicable timeframe (≤7 days from point of diagnosis) to measure specific neuropsychological domains of

delirium (e.g., attention, arousal). As there is evidence to suggest that delirium is associated with functional decline[16], we reviewed studies using repeated tests of functional ability, as well as neuropsychological domains, in tracking delirium recovery in adult patients in acute hospital settings.

METHODS

This systematic review was reported in accordance with the Preferred Reporting of Items in Systematic Reviews and Meta-Analysis (PRISMA) statement (http://prisma-statement.org/documents/PRISMA_2020_checklist.pdf).

Protocol and registration

The study protocol was prospectively registered with PROSPERO (https://www.crd.york.ac.uk/PROSPERO/ registration number CRD42021287331).

Selection criteria

Study inclusion criteria were: (1) hospitalised patients aged ≥18 years diagnosed with delirium according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) or International Statistical Classification of Diseases and Health Related Problems (ICD) criteria, or validated tool, including the Confusion Assessment Method (CAM)[7] or CAM for the Intensive Care Unit (CAM-ICU)[17], Delirium Rating Scale Revised-98 (DRS-R98)[18], Nursing Delirium Screening Scale (Nu-DESC)[19], Memorial Delirium Assessment Scale (MDAS)[20], 4AT[6] and Delirium Index (DI)[21]; (2) baseline data from point of diagnosis of delirium or initial delirium assessment; (3) data from at least one repeated assessment within seven days from baseline (a clinically applicable timeframe; the average delirium episode has been suggested to be ~seven days[22]); studies with further assessments beyond seven days were excluded (4) data collected using neuropsychological assessment tools on at least one symptom domain of delirium (including general cognition, language, attention, level of arousal, memory, orientation, affect and distress), or tools measuring functional ability (e.g., mobility) in the context of repeated assessments. Studies reporting the overall score of a delirium-specific assessment tool (such as the DRS-R98) without reporting change in the individual neuropsychological or functional domains measured by the tool were excluded. Studies of patients without a diagnosis of delirium, or with delirium tremens, were excluded. Studies not written in English were also excluded.

Data sources

An inclusive search strategy was developed with an experienced librarian using selected keywords relating to delirium, key delirium symptom domains, repeated assessment and hospital setting (Appendix 1). The keywords relating to delirium were chosen based on the search syntax published by the National Health and Clinical Excellence (NICE) guidance for delirium[23]. We searched the following databases: MEDLINE® (OVID), EMBASE (OVID), PsycINFO (EBSCO), CINAHL, ClinicalTrials.gov and The Cochrane Central Register of Controlled Trials from 1980 (the year delirium was introduced in DSM-3) to October 14th 2022. An initial search was conducted on October 26th 2021. No restrictions on study design were imposed. We checked reference lists of included articles for further articles of potential relevance. Members of the European Delirium Association, Australasian Delirium Association and American Delirium Society were contacted for eligible studies.

Study selection and data extraction

Two reviewers (HM and EN) independently examined titles and abstracts for eligibility. Eligible studies underwent full text screening independently by both reviewers, resolving any disagreement by discussion with another reviewer (ZT). Data were extracted independently by the two reviewers, with any discrepancies resolved via consensus. Collected data for each study included: study design and setting, population, age, sex, co-morbid illness, number of patients with delirium and dementia, delirium assessment tool(s) used, comparator, change in specific domain(s) of delirium or function (the primary outcome), time between assessment points, statistics used, conclusion of the study and study quality. The main outcome of interest was delirium recovery, specifically a change in a symptom domain of delirium (e.g., attention, arousal, etc.) or a functional domain (e.g., mobility, independence in activities of daily living (ADL), etc.) within seven days from baseline (i.e., first assessment point at which all participants were delirium-positive).

Risk of Bias assessment

Two independent reviewers (HM and EN) assessed risk of bias (RoB) and agreed by consensus using a modified version of the Risk of Bias Assessment Tool for Non-Randomised Studies (RoBANS)[24]. The components of the RoBANS were adapted for the aim of this systematic review, e.g., a study was deemed to be high risk if the authors considered confounding variables in their analyses of group differences (e.g., statistical analyses of difference between intervention and control groups) but not their analyses of change in a specific symptom domain over time (e.g., statistical analyses of change in attention from baseline to follow-up in delirium sample). The RoBANS comprised six components:

the selection of participants, confounding variables, measurement of exposure, the blinding of outcome assessments, incomplete outcome data and selective outcome. The included studies were assessed as high, low and unclear risk based on the modified RoBANS criteria (Appendix 2).

Narrative data synthesis

The possibility of conducting a meta-analysis was explored but ultimately was judged not feasible due to the heterogeneity of the methodology and outcome measures in included studies. As between-study heterogeneity could not be assessed quantitatively, conclusions were based on narrative data synthesis.

RESULTS

Study selection

The search yielded 6,533 papers. From the initial search, 285 were included for full-text screening based on the title and abstract. Fifty-seven papers were not available (authors were contacted where possible) or not published (conference abstracts only, protocols, etc.). From the remaining 228 papers, 189 were excluded as they did not meet the inclusion criteria. The main reasons for exclusion were no actual results (i.e., mean assessment scores) being reported (N = 60), no measurement of specific symptom domains (N = 30), and first repeat assessment being reported at >7 days (N = 14). Many papers had multiple exclusion reasons (N = 57). In total, 39 papers (reporting 32 studies) were included. Figure 1 illustrates the PRISMA flowchart of study selection, including all reasons for exclusion, etc.

[INSERT FIGURE 1]

Study characteristics

A total of 3,434 participants were included in the review population (N range = 10-269, median = 88), of which 2,370 participants had delirium. Nineteen studies included an all-delirious sample. Thirteen studies included a mixed sample of participants with and without delirium. Nine studies included participants with dementia (range = 15-74% of the sample) and 12 studies did not report whether they included dementia patients (including one Boettger et al paper[25] that did not report percentage of dementia patients in their study sub-sample). Although all included studies considered repeat assessment scores separately for delirious subgroups, eight papers did not report descriptive characteristics for participants with and without delirium separately. There was heterogeneity in the

included studies' method of reporting demographic information, preventing overall mean age, range, etc. being reported. The descriptive characteristics for individual studies are presented in Table 1.

Assessment tools

Twenty-one different assessment tools were used in the included studies, each measuring at least one symptom domain (Table 2 and Appendix 3, Figure S1). The three assessment tools most frequently used were the MMSE[14] measuring general cognition (14/32 studies), the DRS-R98[18] measuring delirium presence and severity (9/32 studies), and the Karnofsky Performance Scale (KPS)[26,27] (5/32 studies) measuring functional impairment.

Only four assessments used were delirium-specific: the DRS-R98[18], the Cognitive Test for Delirium (CTD)[28], The Edinburgh Delirium Test Box (EDTB)[29] and the DelApp[30], with less than half of the included studies (14 studies) using these tools. Included studies using the DRS-R98 reported results for its constituent 13 symptom domains e.g., hallucinations, delusions, motor agitation, etc., in addition to the overall DRS-R98 score. For the purpose of this review, the DRS-R98's 'sleep-wake cycle disturbance' item was considered to reflect level of arousal. There were 18 tools not specific to delirium which assessed individual symptom domains in the context of delirium recovery, e.g., Montreal Cognitive Assessment (MoCA)[31], Clock Drawing Test[32], Digit Span Task[10], Cognitive Estimation Task[33], Modified Barthel Index (MBI)[34], etc. (Table 2).

[INSERT TABLES 1-3]

Summary of neuropsychological and functional symptom domains

Included studies assessed a total of 15 neuropsychological or functional symptom domains, including general cognition, attention, level of arousal, psychotic symptoms (e.g., hallucinations and delusions), functional ability, depression and anxiety, and the additional individual DRS-R98 items of lability of affect, language, thought process abnormalities, motor agitation, motor retardation, orientation, short-term memory, long-term memory, and visuospatial ability. The three domains most frequently assessed in the included studies were general cognition (20/32 studies), measured mostly by the MMSE[14], and level of arousal (10/32 studies) and attention (9/32 studies), both primarily measured using the DRS-R98[18].

Longitudinal change in neuropsychological and functional symptom domains

A total of 27 papers analysed longitudinal change in neuropsychological domains or functional ability in participants with delirium, whilst 12 papers solely analysed group differences in deficits in symptom domains e.g., attention in dementia vs non-dementia groups. The results of the individual

papers are reported in Table 3. No studies investigated longitudinal change in symptom domains in the context of delirium recovery.

Of the studies investigating symptom change over time, improvement (within seven days from baseline) was observed in general cognition (12/16 papers reporting improvement in symptoms), functional ability (9/10 papers), psychotic symptoms (4/6 papers), attention (3/5 papers), level of arousal (2/5 papers), motor agitation (2/4 papers), orientation (2/4 papers), visuospatial ability (2/4 papers), lability of affect (2/4 papers), thought process abnormalities (1/4 papers), short-term memory (1/4 papers), language (1/4 papers), and depression and anxiety (1/1 papers). No improvement was found in long-term memory (0/4 papers) or motor retardation (0/4 papers) in the studies analysing longitudinal change.

The symptom domains most frequently assessed for longitudinal change were general cognition, functional ability, attention, level of arousal and psychotic symptoms. Thus, we highlighted the results from the papers reporting these neuropsychological and functional symptom domains in a narrative data synthesis. As few studies analysed change over time (<5 papers per domain) in the remaining 12 domains, it is difficult to draw conclusions on whether scores for these domains improved longitudinally.

General Cognition

Sixteen papers (15 studies) analysed change in general cognition in participants with delirium over time (≤7 days), with 12 papers (11 studies) reporting an improvement[35-46]. Of the two papers (two studies) that analysed longitudinal change but found no significant results[47,48], one had a small sample size (N = 14)[47] and the other did not report the proportion of the sample who had a dementia diagnosis[48]. Two studies reported longitudinal change in general cognition beyond the seven-day timeframe[49,50]. Most studies analysing longitudinal change used the MMSE (12/14)[35,42,48,50]. Five papers (five studies) reported solely group differences but not change over time, comparing delirious and non-delirious groups[51,52], dementia vs non-dementia groups[53], delirium subtypes (i.e., full-syndromal/subsyndromal delirium and resolving/persistent delirium)[54] or treatment and control groups for drug interventions[55].

Functional Ability

Ten papers (five studies)[25, 35, 40, 42, 56-61] analysed functional recovery over time in the context of delirium, with nine (four studies) observing improvement. Boettger et al.[25,56-61] analysed change in KPS scores over time and found an improvement in functional performance. One study[40] found no significant change, however, its sample had a high mean age (84 years) and a large proportion of participants (74%) had dementia. No studies investigated group differences in functional performance.

Attention

Five papers (five studies)[35,41,50,62,63] analysed whether attention changed over time. Of those five studies, three reported improvements[35,62,63]. Although the remaining two studies did investigate longitudinal change for performance on attention tasks, they did not report analyses conducted on scores within the seven-day timeframe[41,50]. Assessments used by the included studies to measure attention varied (e.g., EDTB, DelApp) but the DRS-R98 attention sub-item was used most frequently (3/5 papers). Three additional papers reported group differences in attention deficits, comparing performance on assessments (EDTB-ICU, DelApp-ICU and DRS-R98) between delirious and non-delirious patients or between delirium subtypes (e.g., full syndromal vs subsyndromal)[29,30,54].

Arousal

Five papers (five studies)[50,62-65] analysed whether level of arousal changed (i.e., closer to 'normal' arousal; 'alert', 'calm', etc.) over time. Only two of these reported improvement in level of arousal[63,64], whilst one study found no improvement[62]. This study[62] had a small sample included in their analyses (N = 18) and arousal was only measured by the DRS-R98 item of 'sleep-wake cycle disturbance'. The remaining two studies[50,65] investigated longitudinal change in level of arousal, but they did not report analyses for results within the seven-day timeframe. Three studies investigated group differences in arousal, comparing the effectiveness of drug interventions or performance between delirium subtypes[54,66,67].

Psychotic features

Five papers (five studies)[39,43,44,62,63] analysed longitudinal change in psychotic features. Four of these studies reported an improvement in psychotic features over time[39,43,44,63]. The remaining paper[62] found no effects for improvement on DRS-R98 items 'hallucinations' or 'delusions' within the seven-day timeframe, however, results for only 18/25 delirious participants were reported due to patient discharge before day four. An additional paper[54] only conducted analyses for group differences, comparing presentation of psychotic features between delirium subtypes; longitudinal change in scores were presented solely in graph format.

[INSERT FIGURE 2]

Risk of Bias

There was variability in RoB between included studies, however, the overall RoB was moderate to high (Appendix 4). The RoBANS component with the lowest risk was selection of participants. All

papers presented a low risk in selection of participants as they clearly stated broad participant eligibility criteria, except two papers which presented an unclear risk[45,47]. All other criteria presented a high or unclear risk. In measurement of exposure, 26 papers were low risk whilst 13 papers were either unclear or high risk due to missing key information on how assessments were carried out and data were collected[25,35,38,42,55,56,58,59,62,63,66,68,69]. Only 10 papers were low risk in blinding of outcome measures[5,20,44,46,48,51,55,66,67,70], with most papers presenting unclear risk due to a lack of clarity around blinding details. In terms of incomplete outcome data, 12 papers had high risk due to an absence of a published protocol[5,29,30,43,45,46,48,55,61,65,67,68]. The highest RoB was found regarding confounding variables; only five papers [29,39,45,54,57,61] had low RoB in confounding variables, with the remaining papers presenting a high risk. Many studies were deemed to be high risk if they considered confounding variables (e.g., in analyses of group differences) but not specifically in the analyses of longitudinal change, i.e., they did not consider variables such as age, dementia diagnosis, serious illness, etc. in the improvement of symptom domains over repeated assessments. The individual RoB ratings for each study are presented in Appendix 4.

DISCUSSION

This systematic review comprehensively outlines the assessment tools currently used in research settings to measure change in neuropsychological and functional symptom domains of delirium in the context of delirium recovery. The significant methodological heterogeneity in the included studies highlights the lack of a standardised method for measuring change in delirium symptoms over repeat assessments.

Evaluation of main findings

Assessment tools

In total, 21 assessment tools were used to measure recovery in 15 neuropsychological or functional domains. The domain most investigated in the included studies was general cognition, mostly assessed by the MMSE. The DRS-R98 was used in 25% of the included studies.

Recovery in neuropsychological and functional domains

Although the included studies generally did not focus on delirium recovery, 27 papers (21 studies) investigated longitudinal change in symptom domains. Of the 27 papers reporting longitudinal change, 24 papers (89%) reported an improvement in one or more domains, except for memory impairment and motor retardation. The symptom domains most frequently assessed for change over time were general cognition, functional ability, attention, level of arousal and psychotic symptoms.

Despite most studies reporting an improvement in general cognition, two studies investigating change in general cognition found no improvement in reported deficits over time[47,48]. This may be due to Lowery et al[47] reporting on a small sample of patients with delirium in a non-representative setting (elective orthopaedic ward). Additionally, Maneeton et al[48] did not report the percentage of participants with a dementia diagnosis. This lack of consideration of the confounding variable of general cognition may have affected their results.

Dementia diagnosis may have affected functional ability scores over time in another study[40]. This study had the largest proportion of participants with delirium superimposed on dementia (74.4%) in their sample. Dementia causes a progressive decline in general cognition[710]and this is directly related to poorer functional ability over time[72]. Delirious patients with dementia may already have had a poor functional baseline at the onset of their delirium and thus may be less likely to show measurable recovery in this domain.

Level of arousal was one of the most assessed symptom domains across included studies; however, only two studies found an improvement ≤7 days[63,64]. One study found that level of arousal did not improve over time[62]; this was likely due to (a) small sample size (N = 18) and (b) the use of DRS-R98 item 'sleep-wake cycle disturbance' to assess arousal, which is a three-point scale that is likely less sensitive to recovery-related changes compared to arousal-specific tool such as the RASS or OSLA.

Most studies reporting longitudinal change in psychotic features found a significant improvement in symptoms over time. Only one study did not find a change[62], which may again be due to the limitations of using only the simple DRS-R98 items 'hallucinations' and 'delusions'. Moreover, psychotic symptoms may be more prominent in hyperactive compared to hypoactive delirious participants[73]; however, this paper with a small sample size (N = 18) did not differentiate between motor subtypes (though hyperactive patients do not necessarily experience more psychotic symptoms than other subtypes[74]), which may have affected the recovery trajectory observed in this study.

These findings point to the potential value of quantitative measurement of neuropsychological and functional domains in assessing delirium recovery. As recently argued, because of the heterogeneity of delirium between individuals, research-grade delirium assessment will need to include measurement of multiple domains using validated tests that are sensitive to change[75]. Reliance on single linear scales is not likely to yield accurate enough information, and capture of delirium recovery across a range of cognitive, affective, perceptual and functional domains will be necessary. Additionally, there are features in which measurements suitable both accurate symptom ascertainment and tracking change over time still need to be developed, such distress.

Strengths and limitations

This is the first systematic literature review evaluating studies which used assessment tools of neuropsychological domains and functional recovery in patients with delirium in the context of recovery. We comprehensively searched databases using inclusive search strategies, constructed with the help of an experienced librarian, and contacted delirium experts and researchers for studies meeting the inclusion criteria. Despite this, we were unable to access 57 studies identified in the title and abstract screening (many due to a lack of publication).

Included studies had high or unclear risk of bias for at least two domains and were of mixed methodological quality overall. There were many studies that did not consider confounding variables such as age, dementia diagnosis, etc. in their analyses of change in symptom domains over time. Recovery in neuropsychological domains may be affected by the participants' baseline cognition, which was not considered in many of the included studies. This is also a general limitation of the neuropsychological testing approach in assessing recovery from delirium.

Moreover, none of the assessment tools used in the included studies are specifically validated for repeat use in delirious patients; it is therefore unclear whether they retain their diagnostic accuracy after baseline (i.e., detection of delirium). Most assessment tools were not delirium-specific and therefore may not be as effective in capturing deficits in individual symptom domains in delirious patients or may not be sensitive enough to subtle changes in delirium symptom presentations. The DRS-R98 was the main delirium-specific tool used. However, it is biased towards hyperactive features of delirium with those in lower arousal states not adequately assessed for deficits in some symptom domains, i.e., patients could score zero (no impairment) if they are unable to be assessed due to their low level of arousal.

A meta-analysis could not be conducted due to considerable heterogeneity in the methodologies used. Studies rarely aimed to specifically investigate recovery as measured by individual delirium symptom domains. Instead, we provide a narrative data synthesis to outline the key symptom domains that improved over time and draw conclusions based on this. As our inclusion criteria did not exclude studies involving interventions, some studies reported change within symptom domains as a result of drug intervention e.g., aripiprazole, haloperidol[25,56-60]. This may have affected the recovery trajectories of symptom domains, e.g., functional recovery in the Boettger et al. study[25,56-60]. In these instances, conclusions of symptom recovery may not be entirely generalisable to an acute hospital delirious population, as although haloperidol in small doses is administered in clinical settings to manage severe symptoms (e.g., distress and agitation), use of aripiprazole is not common practice (in the UK), nor is it advised by the NICE guidelines[23].

Monitoring delirium recovery in clinical practice

This review provides evidence that in research studies delirium and its individual symptom presentations are often monitored on a repeat basis to measure improvement in patients with delirium. Despite these studies suggesting the importance of monitoring delirium recovery, delirium can still be under-detected or under-reported[76] This can result in patients being discharged without documented resolution of delirious symptoms[77,78] Delirium documentation remains inadequate, despite an increase in delirium mentioned in discharge summaries[79]. This may partly be due to a lack of consensus on how delirium should be monitored on a repeat basis. The NICE[23] standards recommend daily observations of a patient's behaviour to identify fluctuations of delirium, however, no validated method for repeated assessment exists. Guidelines on how delirium and symptom domains should be measured for recovery may result in higher quality delirium documentation as well as reduced risk of discharge with active delirium.

Although the MMSE was the most frequently reported tool across included studies, use of the MMSE is not recommended by guidelines for assessment of delirium. In clinical practice, general cognition in patients with delirium is measured more frequently by the Montreal Cognitive Assessment (MoCA)[31]. This tool is recommended for use as a measure of baseline cognition in patients with delirium by the Scottish Intercollegiate Guidelines Network (SIGN)[80], however, only one included study utilised the MoCA[68]. Similarly, the DRS-R98 was widely used as a tool to assess neuropsychological domains, which is also not recommended in national guidelines[23,80] for use in acute clinical practice, possibly due to the long duration of administration. Any method validated for assessment of delirium recovery should be appropriate for use in acute settings, which require efficient and feasible administration at the bedside.

In clinical practice, cases of hyperactive delirium (characterised by agitation, hypervigilance, etc.) are often prioritised over those of hypoactive delirium (characterised by reduced arousal, apathy, etc.) in detection[81,82]. Hypoactive patients are at risk of being undiagnosed due to not engaging with cognitive testing and hence being deemed 'unable to assess'. Any favoured detection of hyperactive over hypoactive delirium could present clinical issues for recovery, as there may be less focus on care for hypoactive patients. As many studies investigated change in level of arousal but few found improvement, future studies should investigate how this symptom domain fluctuates and improves over short, clinically applicable timeframes.

Implications for research and clinical practice

This review has identified an important gap in the field as none of the delirium tools currently in use are specifically tailored to measuring recovery. Future research should address this by focusing on repeat assessments of delirium using delirium-specific tools to track recovery and validating methods

for doing so. Research should also investigate how different symptom domains change in the context of delirium recovery, for instance, whether deficits persist in some symptom domains even after recovery in others. This will be helpful in identifying the symptom domains that are most useful in determining recovery from delirium.

To improve patient care and outcomes, there is a clear need for validated and pragmatic assessments of delirium recovery. This is essential to communicate details of care to the patient and their relatives, to ensure fewer patient discharges with unresolved delirium and to inform treatment. A recent international survey of clinicians involved in delirium care found that the most used repeat assessments measured a range of delirium features used including level of arousal, inattention, motor disturbance and psychotic features[83]. Notably, delirium-specific tools were reported to be used by clinicians rather than tests of general cognition such as the MMSE; this provides an interesting contrast to the research literature summarised in the present review. An existing, validated screening tool with a short administration time, such as the 4AT, may be a pragmatic choice for measuring recovery from delirium via repeat assessments in an acute care setting. Validation studies of this, and other tools, in clinical settings with relevant inclusion criteria and data collection to minimise confounding variables, are now required.

Conclusions

This review identifies a clear gap in the field, highlighting the need for further research on how symptom domains of delirium change in line with recovery and consensus on assessment tools to monitor for recovery. This also provides a rationale for validation of assessment tools for delirium recovery to be pursued in future research.

REFERENCES

1. Wilson JE, Mart MF, Cunningham C, et al. Delirium. *Nature Reviews Disease Primers*. 2020;6(1):1-26. doi:10.1038/s41572-020-00223-4
2. Reynish EL, Hapca SM, De Souza N, Cvoro V, Donnan PT, Guthrie B. Epidemiology and outcomes of people with dementia, delirium, and unspecified cognitive impairment in the general hospital: prospective cohort study of 10,014 admissions. *BMC Medicine*. 2017;15(1):140-140. doi:10.1186/s12916-017-0899-0
3. Kiely DK, Jones RN, Bergmann MA, Murphy KM, Orav EJ, Marcantonio ER. Association between delirium resolution and functional recovery among newly admitted postacute facility patients. *The Journals of Gerontology: Series A*. 2006;61(2):204-208. doi:10.1093/gerona/61.2.204
4. Bickel H, Gradinger R, Kochs E, Förstl H. High risk of cognitive and functional decline after postoperative delirium. *Dementia and Geriatric Cognitive Disorders*. 2008;26(1):26-31. doi:10.1159/000140804
5. Cole MG, McCusker J, Bailey R, et al. Partial and no recovery from delirium after hospital discharge predict increased adverse events. *Age and Ageing*. 2017;46(1):90-95. doi:10.1093/ageing/afw153
6. Shenkin SD, Fox C, Godfrey M, 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 Medicine*. 2019;17(1):1-14. doi:10.1186/s12916-019-1367-9
7. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Annals of Internal Medicine*. 1990;113(12):941-948. doi:10.7326/0003-4819-113-12-941
8. 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. *Palliative Medicine*. 2010;24(6):561-565. doi:10.1177/0269216310371556
9. Regan NA, Ryan DJ, Boland E, et al. Attention! A good bedside test for delirium? *Journal of Neurology, Neurosurgery & Psychiatry*. 2014;85(10):1122. doi:10.1136/jnnp-2013-307053
10. Groth-Marnat G, Baker S. Digit Span as a measure of everyday attention: a study of ecological validity. *Perceptual and Motor Skills*. 2003;97(3):1209-1218. doi:10.2466/pms.2003.97.3f.1209
11. Tieges Z, Stott DJ, Shaw R, et al. A smartphone-based test for the assessment of attention deficits in delirium: A case-control diagnostic test accuracy study in older hospitalised patients. *PloS One*. 2020;15(1):e0227471. doi:10.1371/journal.pone.0227471

12. Sessler CN, Gosnell MS, Grap MJ, et al. The Richmond Agitation–Sedation Scale: validity and reliability in adult intensive care unit patients. *American Journal of Respiratory and Critical Care Medicine*. 2002;166(10):1338-1344. doi:10.1164/rccm.2107138
13. Hall R, Stíobhairt 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. *International Journal of Geriatric Psychiatry*. 2020;35(9):1021-1027. doi:10.1002/gps.5324
14. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6
15. Adamis D, Devaney A, Shanahan E, McCarthy G, Meagher D. Defining 'recovery' for delirium research: a systematic review. *Age and Ageing*. 2014;44(2):318-321. doi:10.1093/ageing/afu152
16. Balas MC, Happ MB, Yang W, Chelluri L, Richmond T. Outcomes associated with delirium in older patients in surgical ICUs. *Chest*. 2009;135(1):18-25. doi:10.1378/chest.08-1456
17. Ely EW, Inouye SK, Bernard GR, 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
18. Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N. Validation of the Delirium Rating Scale-Revised-98. *The Journal of Neuropsychiatry and Clinical Neurosciences*. 2001;13(2):229-242. doi:10.1176/jnp.13.2.229
19. Hargrave A, Bourgeois J, Bastiaens J, et al. Validation of a nurse-based delirium screening tool for hospitalized patients. *Neurology*. 2016;86(16):P6.220.
20. Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K, Passik S. The Memorial Delirium Assessment Scale. *Journal of Pain and Symptom Management*. 1997;13(3):128-137. doi:10.1016/S0885-3924(96)00316-8
21. McCusker J, Cole MG, Dendukuri N, Belzile E. The Delirium Index, a measure of the severity of delirium: new findings on reliability, validity, and responsiveness. *Journal of the American Geriatrics Society*. 2004;52(10):1744-1749. doi:10.1111/j.1532-5415.2004.52471.x
22. Nguyen TH, Atayee RS, Derry KL, Hirst J, Biondo A, Edmonds KP. Characteristics of hospitalized patients screening positive for delirium. *American Journal of Hospice and Palliative Medicine®*. 2019;37(2):142-148. doi:10.1177/1049909119867046
23. National Institute for Health and Care Excellence (NICE). Delirium: Prevention, Diagnosis and Management Accessed August 2, 2022. https://www.nice.org.uk/guidance/cg103
24. Kim SY, Park JE, Lee YJ, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*. 2013;66(4):408-414. doi:10.1016/j.jclinepi.2012.09.016

25. Boettger S, Friedlander M, Breitbart W, Passik S. Aripiprazole and haloperidol in the treatment of delirium. *Australian & New Zealand Journal of Psychiatry*. 2011;45(6):477-482. doi:10.3109/00048674.2011.543411
26. Karnofsky DA, Abelmann WH, Craver LF, Burchenal JH. The use of the nitrogen mustards in the palliative treatment of carcinoma: with particular reference to bronchogenic carcinoma. *Cancer*. 1948;1(4):634-656. doi:10.1002/1097-0142(194811)1:4<634::AID-CNCR2820010410>3.0.CO;2-L
27. Mor V, Laliberte L, Morris JN, Wiemann M. The Karnofsky Performance Status Scale: an examination of its reliability and validity in a research setting. *Cancer*. 1984;53(9):2002-2007. doi:10.1002/1097-0142(19840501)53:9<2002::AID-CNCR2820530933>3.0.CO;2-W
28. Hart RP, Levenson JL, Sessler CN, Best AM, Schwartz SM, Rutherford LE. Validation of a Cognitive Test for Delirium in Medical ICU Patients. *Psychosomatics*. 1996;37(6):533-546. doi:10.1016/S0033-3182(96)71517-7
29. Green C, Hendry K, Wilson ES, et al. A novel computerized test for detecting and monitoring visual attentional deficits and delirium in the ICU. *Crit Care Med*. 2017;45(7):1224-1231. doi:10.1097/ccm.0000000000002477
30. Tang E, Laverty M, Weir A, et al. Development and feasibility of a smartphone-based test for the objective detection and monitoring of attention impairments in delirium in the ICU. *Journal of Critical Care*. 2018;48:104-111. doi:10.1016/j.jcrc.2018.08.019
31. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*. 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x
32. Adamis D, Morrison C, Treloar A, Macdonald AJD, Martin FC. The performance of the Clock Drawing Test in elderly medical inpatients: does it have utility in the identification of delirium? *Journal of Geriatric Psychiatry and Neurology*. 2005;18(3):129-133. doi:10.1177/0891988705277535
33. Shallice T, Evans ME. The involvement of the frontal lobes in cognitive estimation. *Cortex*. 1978;14(2):294-303. doi:10.1016/S0010-9452(78)80055-0
34. Mahoney F, Barthel D. Functional evaluation: the Barthel Index. *Maryland State Medical Journal*. 1965;14(2):61-65. doi:10.1037/t02366-000
35. Bellelli G, Speciale S, Morghen S, Torpilliesi T, Turco R, Trabucchi M. Are fluctuations in motor performance a diagnostic sign of delirium? *Journal of the American Medical Directors Association*. 2011;12(8):578-583. doi:10.1016/j.jamda.2010.04.010
36. Breitbart W, Marotta R, Platt MM, et al. A double-blind trial of haloperidol, chlorpromazine, and lorazepam in the treatment of delirium in hospitalized AIDS patients. *FOCUS*. 2005;3(2):333-340. doi:10.1176/foc.3.2.333

37. Chong E, Tay L, Chong MS. Identifying phenomenological differences and recovery of cognitive and non-cognitive symptomatology among delirium superimposed upon dementia patients (DsD) versus those without dementia (DaD) in an acute geriatric care setting. *International Psychogeriatrics*. 2015;27(10):1695-1705. doi:10.1017/S1041610215000770
38. Gagnon B, Low G, Schreier G. Methylphenidate hydrochloride improves cognitive function in patients with advanced cancer and hypoactive delirium: a prospective clinical study. *Journal of Psychiatry and Neuroscience*. 2005;30(2):100.
39. Kishi Y, Kato M, Okuyama T, Thurber S. Treatment of delirium with risperidone in cancer patients. *Psychiatry and Clinical Neurosciences*. 2012;66(5):411-417. doi:10.1111/j.1440-1819.2012.02346.x
40. Lam C, Tay L, Chan M, Ding YY, Chong MS. Prospective observational study of delirium recovery trajectories and associated short-term outcomes in older adults admitted to a specialized delirium unit. *Journal of the American Geriatrics Society*. 2014;62(9):1649-1657. doi:10.1111/jgs.12995
41. Lou MF, Dai YT, Huang GS, Yu PJ. Postoperative cognitive changes among older Taiwanese patients. *Journal of Clinical Nursing*. 2003;12(4):579-588. doi:10.1046/j.1365-2702.2003.00753.x
42. Mittal D, Jimerson NA, Neely EP, et al. Risperidone in the treatment of delirium: results from a prospective open-label trial. *Journal of Clinical Psychiatry*. 2004;65:662-667. doi:10.4088/JCP.v65n0510
43. Parellada E, Baeza I, de Pablo J, Martínez G. Risperidone in the treatment of patients with delirium. *Journal of Clinical Psychiatry*. 2004;65(3):348-353. doi:10.4088/JCP.v65n0310
44. Pintor L, Fuente E, Bailles E, Matrai S. Study on the efficacy and tolerability of amisulpride in medical/surgical inpatients with delirium admitted to a general hospital. *European Psychiatry*. 2009;24(7):450-455. doi:10.1016/j.eurpsy.2009.05.013
45. Saczynski JS, Marcantonio ER, Quach L, et al. Cognitive trajectories after postoperative delirium. *New England Journal of Medicine*. 2012;367(1):30-39. doi:10.1056/NEJMoa1112923
46. Yoon HJ, Park KM, Choi WJ, et al. Efficacy and safety of haloperidol versus atypical antipsychotic medications in the treatment of delirium. *BMC Psychiatry*. 2013;13(1):240. doi:10.1186/1471-244X-13-240
47. Lowery DP, Wesnes K, Brewster N, Ballard C. Quantifying the association between computerised measures of attention and confusion assessment method defined delirium: a prospective study of older orthopaedic surgical patients, free of dementia. *International Journal of Geriatric Psychiatry*. 2008;23(12):1253-1260. doi:10.1002/gps.2059

48. Maneeton B, Maneeton N, Srisurapanont M, Chittawatanarat K. Quetiapine versus haloperidol in the treatment of delirium: a double-blind, randomized, controlled trial. *Drug Design, Development and Therapy*. 2013;7:657. doi:10.2147/DDDT.S45575
49. Milisen K, Foreman MD, Abraham IL, et al. A nurse-led interdisciplinary intervention program for delirium in elderly hip-fracture patients. *Journal of the American Geriatrics Society*. 2001;49(5):523-532. doi:10.1046/j.1532-5415.2001.49109.x
50. Leonard M, Adamis D, Saunders J, Trzepacz P, Meagher D. A longitudinal study of delirium phenomenology indicates widespread neural dysfunction. *Palliative and Supportive Care*. 2015;13(2):187-196. doi:10.1017/S147895151300093X
51. Deschodt M, Braes T, Flamaing J, et al. Preventing delirium in older adults with recent hip fracture through multidisciplinary geriatric consultation. *Journal of the American Geriatrics Society*. 2012;60(4):733-739. doi:10.1111/j.1532-5415.2012.03899.x
52. Cole M, Primeau F, Bailey R, et al. Systematic intervention for elderly inpatients with delirium: a randomized trial. *CMAJ: Canadian Medical Association Journal*. 1994;151(7):965.
53. Lingehall HC, Smulter NS, Lindahl E, et al. Preoperative cognitive performance and postoperative delirium are independently associated with future dementia in older people who have undergone cardiac surgery: a longitudinal cohort study. *Critical Care Medicine*. 2017;45(8):1295. doi:10.1097/CCM.0000000000002483
54. Meagher D, Adamis D, Trzepacz P, Leonard M. Features of subsyndromal and persistent delirium. *British Journal of Psychiatry*. 2012;200(1):37-44. doi:10.1192/bjp.bp.111.095273
55. Tahir TA, Eeles E, Karapareddy V, et al. A randomized controlled trial of quetiapine versus placebo in the treatment of delirium. *Journal of Psychosomatic Research*. 2010;69(5):485-490. doi:10.1016/j.jpsychores.2010.05.006
56. Boettger S, Breitbart W. An open trial of aripiprazole for the treatment of delirium in hospitalized cancer patients. *Palliative & Supportive care*. 2011;9(4):351-357. doi:10.1017/S1478951511000368
57. Boettger S, Breitbart W, Jenewein J, Boettger S. Delirium and functionality: the impact of delirium on the level of functioning. *The European Journal of Psychiatry*. 2014;28(2):86-95. doi:10.4321/S0213-61632014000200002
58. Boettger S, Jenewein J, Breitbart W. Delirium in advanced age and dementia: a prolonged refractory course of delirium and lower functional status. *Palliative and Supportive Care*. 2015;13(4):1113-1121. doi:10.1017/S1478951514000972
59. Boettger S, Jenewein J, Breitbart W. Delirium management in patients with cancer: dosing of antipsychotics in the delirium subtypes and response to psychopharmacological management. *German Journal of Psychiatry*. 2014;17(1):10-18. doi:10.5167/uzh-107992

60. Boettger S, Passik S, Breitbart W. Treatment characteristics of delirium superimposed on dementia. *International Psychogeriatrics*. 2011;23(10):1671-1676. doi:10.1017/S1041610211000998
61. Flaherty JH, Steele DK, Chibnall JT, Vasudevan VN, Bassil N, Vegi S. An ACE unit with a delirium room may improve function and equalize length of stay among older delirious medical inpatients. *The Journals of Gerontology: Series A*. 2010;65A(12):1387-1392. doi:10.1093/gerona/glq136
62. Maneewong J, Maneeton B, Maneeton N, et al. Delirium after a traumatic brain injury: predictors and symptom patterns. *Neuropsychiatric Disease and Treatment*. 2017;13:459. doi:10.2147/NDT.S128138
63. Matsuda Y, Nakao Y, Yabe M, Tsuruta R, Takemura M, Inoue K. Association of the clinical subtype and etiology for delirium with the outcome after risperidone monotherapy in patients having cancer. *Osaka City Med J*. 2016;62:19-28.
64. Kim SY, Kim JM, Kim SW, et al. Do the phenotypes of symptom fluctuation differ among motor subtypes in patients with delirium? *Journal of Pain and Symptom Management*. 2018;56(5):667-677. doi:10.1016/j.jpainsymman.2018.07.022
65. Lagarto L, Albuquerque E, Loureiro D, et al. Arousal changes and delirium in acute medically-ill male older patients with and without dementia: a prospective study during hospitalization. *Aging & Mental Health*. 2020;24(5):820-827. doi:10.1080/13607863.2018.1548569
66. Liu L, Yuan Q, Wang Y, et al. Effects of dexmedetomidine combined with sufentanil on postoperative delirium in young patients after general anesthesia. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*. 2018;24:8925. doi:10.12659/MSM.911366
67. Hui D, Frisbee-Hume S, Wilson A, et al. Effect of lorazepam with haloperidol vs haloperidol alone on agitated delirium in patients with advanced cancer receiving palliative care: a randomized clinical trial. *JAMA*. 2017;318(11):1047-1056. doi:10.1001/jama.2017.11468
68. Williams J, Finn K, Melvin V, Meagher D, McCarthy G, Adamis D. The association of serum levels of brain-derived neurotrophic factor with the occurrence of and recovery from delirium in older medical inpatients. *BioMed Research International*. 2017;2017:5271395. doi:10.1155/2017/5271395
69. Mercadante, S., Adile, C., Ferrera, P., Cortegiani, A., & Casuccio, A. Symptom expression in patients with advanced cancer admitted to an acute supportive/palliative care unit with and without delirium. *The Oncologist*. 2019;24(6):358-364. doi: 10.1634/theoncologist.2018-0244
70. Yang, J., Choi, W., Ko, Y. H., Joe, S. H., Han, C., & Kim, Y. K. Bright light therapy as an adjunctive treatment with risperidone in patients with delirium: a randomized, open, parallel

group study. *General hospital psychiatry*. 2012;34(5): 546-551. doi: 10.1016/j.genhosppsych.2012.05.003

71. Ballard C, O'Brien J, Morris CM, et al. The progression of cognitive impairment in dementia with Lewy bodies, vascular dementia and Alzheimer's disease. *International Journal of Geriatric Psychiatry*. 2001;16(5):499-503. doi:10.1002/gps.381
72. Liu-Seifert H, Siemers E, Sundell K, et al. Cognitive and functional decline and their relationship in patients with mild Alzheimer's dementia. *Journal of Alzheimer's Disease*. 2015;43:949-955. doi:10.3233/JAD-140792
73. Boettger S, Breitbart W. Phenomenology of the subtypes of delirium: phenomenological differences between hyperactive and hypoactive delirium. *Palliative and Supportive Care*. 2011;9(2):129-135. doi:10.1017/S1478951510000672
74. Stagno D, Gibson C, Breitbart W. The delirium subtypes: A review of prevalence, phenomenology, pathophysiology, and treatment response. *Palliative and Supportive Care*. 2004;2(2):171-179. doi:10.1017/S1478951504040234
75. MacLullich, A. M., Hosie, A., Tieges, Z., & Davis, D. H. Three key areas in progressing delirium practice and knowledge: recognition and relief of distress, new directions in delirium epidemiology and developing better research assessments. *Age and Ageing*, 2022;51(11):271. doi: 10.1093/ageing/afac271
76. Collins N, Blanchard MR, Tookman A, Sampson EL. Detection of delirium in the acute hospital. *Age and Ageing*. 2010;39(1):131-135. doi:10.1093/ageing/afp201
77. van Zyl LT, Davidson PR. Delirium in hospital: an underreported event at discharge. *The Canadian Journal of Psychiatry*. 2003;48(8):555-560. doi:10.1177/070674370304800807
78. Kakuma R, Du Fort GG, Arsenault L, et al. Delirium in older emergency department patients discharged home: effect on survival. *Journal of the American Geriatrics Society*. 2003;51(4):443-450. doi:10.1046/j.1532-5415.2003.51151.x
79. Chuen VL, Chan ACH, Ma J, Alibhai SMH, Chau V. The frequency and quality of delirium documentation in discharge summaries. *BMC Geriatrics*. 2021;21(1):307. doi:10.1186/s12877-021-02245-3
80. Scottish Intercollegiate Guidelines Network (SIGN). Risk Reduction and Management of Delirium. Accessed August 2, 2022. https://www.sign.ac.uk/
81. Meagher D. Motor subtypes of delirium: Past, present and future. *International Review of Psychiatry*. 2009;21(1):59-73. doi:10.1080/09540260802675460
82. van Velthuijsen EL, Zwakhalen SMG, Mulder WJ, Verhey FRJ, Kempen GIJM. Detection and management of hyperactive and hypoactive delirium in older patients during hospitalization: a retrospective cohort study evaluating daily practice. *International Journal of Geriatric Psychiatry*. 2018;33(11):1521-1529. doi:10.1002/gps.4690

83. Noble, E., McCartney, H., MacLullich, A. M., Shenkin, S. D., Muniz-Terrera, G., Evans, J. J., ... & Tieges, Z. Assessing recovery from delirium: an international survey of healthcare professionals involved in delirium care. *Delirium Communications*. 2019

Table 1: Descriptive characteristics of included studies

Study ID	Study design	Setting	Delirium	Dementia (% of sample with delirium)	Mean age (SD)	Sex (female, male)	Delirium diagnostic tool	Repeat assessment tool	Symptom domains	Number of assess-ments
GENERAL COGNITION (and other domains inc. functional ability, psychotic symptoms, etc.)										
Bellelli et al (2011)	Prospective case-controlled study	Rehabilitation and aged care facility	30	15 (50%)	81.4 (6.2)	43, 17	CAM	TCT, Tinetti Scale, Clock Drawing Test, Digit Span Task, Cognitive Estimation Test, Verbal Abstraction task	General cognition, functional ability	2
Breitbart et al (1996)	Randomised comparison trial	Hospital	30	Not reported	39.2 (SD not reported)	7, 23	DSM-3-R and DRS-R98	MMSE	General cognition	3
Chong et al (2015)†	Prospective cohort study	Geriatric monitoring unit	234	174 (74.4%)	84.1 (7.4)	132, 102	CAM	MMSE and DRS-R98	General cognition, DRS-R98 cognitive sub-score	5
Cole et al (1994)	Randomised controlled trial	Primary acute care hospital	88	22 (25.0%)	86.1 (SD not reported)	57, 31	DSM-3-R	SPMSQ and CGBRS	General cognition, Abnormal behaviour/ADLs	2
Deschodt et al (2012)	Non-randomised (parallel group) controlled trial	Hospital (trauma)	40	34 (19.9%)¶	80.8 (SD not reported)§	109, 62¶	CAM and Delirium Index	MMSE	General cognition	3
Gagnon et al (2005)	Prospective clinical case-control study	Hospital, palliative care unit	14	Not reported	41-80 (mean and SD not reported)	5, 9	DSM-4	MMSE	General cognition	3

Kishi et al (2012)	Prospective study repeated measures	Hospital (psychiatric service)	29	0	68.9 (SD not reported)	9, 20	DSM-4-TR and DRS-R98	MMSE and DRS-R98	General cognition, DRS-R98 sub-items	2
Lam et al (2014)†	Prospective observational study	Hospital, geriatric monitoring unit	234	174 (74.4%)	84.1 (7.4)	132, 102	CAM	MMSE, Modified Barthel Index	General cognition, functional ability	4
Leonard et al (2013) ‡	Prospective observational study	Hospice, palliative care unit	100	27 (27.0%)	70.2 (10.5)	49, 51	DSM-4 and CAM	CTD and DRS-R98	General cognition, DRS-R98 sub-items	2
Lingehall et al (2017)	Prospective cohort study	Hospital (cardiothoracic)	64	0	76.5 (SD not reported)¶	35, 79¶	DSM-4-TR	MMSE	General cognition	2
Lou et al (2003)	Prospective cohort study	Hospital (acute)	41	Not reported	73.9 (6.5)¶	52, 54¶	DSM-4	MMSE	General cognition, MMSE sub-items	3
Lowery et al (2008)	Prospective cohort study	Hospital, orthopaedic surgery	14	0	76.5 (4.5)¶	53, 41¶	CAM	MMSE	General cognition	3
Maneeton et al (2013)	Randomised controlled trial	Hospital, tertiary care setting	52	Not reported	56.8 (11.8)	17, 35	DSM-4-TR and CAM	DRS-R98	General cognition (DRS-R98 cognitive sub-score)	8

Meagher et al (2012) ‡	Prospective observational study	Palliative care	100	27 (27.0%)	70.2 (10.5)	49, 51	DSM-4 and CAM	DRS-R98 and CTD	General cognition, DRS-R98 sub-items	2
Milisen et al (2001)	Prospective before-after design (sequential) study	Emergency room and traumatologic-al units, academic medical centre	26	18 (15.0%) ¶	80 (median control), 82 (median interventio n) (mean and SD not reported)	97, 23¶	CAM	MMSE	General cognition	3
Mittal et al (2004)	Prospective case-controlled clinical study	Hospital (general medical/surgi-cal)	10	0	64.7 (4.8)	2, 8	CAM, DSM-4 and DRS-R98	CTD and KPS	General cognition and functional ability	6
Parellada et al (2004)	Prospective observational study	Hospital	64	Not reported	67.3 (11.4)	24, 40	DSM-4	MMSE and PANSS-P	General cognition, Psychotic symptoms	8
Pintor et al (2009)	Prospective clinical study	Tertiary care, general hospital	31	0	73 (15.7) ¶	19, 21¶	DSM-4	MMSE and PANSS-P	General cognition, Psychotic symptoms	7
Saczynski et al (2012)	Prospective observational study (cohort)	Hospital	103	Not reported	75 (6.5)	30, 73	CAM	MMSE, digit span test	General cognition, attention	2
Tahir et al (2010)	Randomised controlled trial	Hospital (medical,	42	0	84.2 (8.3)	30, 12	DSM-4 and DRS-R98	MMSE and DRS-R98	General cognition, DRS-R98	5

		surgical and orthopaedics)							cognitive sub-score	
William et al (2017)	Prospective observational study	Hospital (acute)	58	86 (43.4%) ¶	80.6 (6.8)	92, 106¶	DRS-R98	MOCA	General cognition	2
Yoon et al (2013)	Prospective clinical study	Hospital, tertiary care setting	80	0	71.8 (11.5)	44, 36	DSM-4-TR	MMSE and DRS-R98	General cognition, DRS-R98 cognitive sub-score	4
INDIVIDUAL NEUROPSYCHOLOGICAL DOMAINS										
Green et al (2017)	Prospective case-controlled study	Hospital (ICU)	15	0	65 (SD not reported)	Not reported	CAM-ICU	EDTB-ICU and RASS	Attention, arousal/agitation	5
Hui et al (2017)	Randomised controlled trial	Acute palliative care unit	58	0	65.0 (SD not reported)	26, 32	DSM-4-TR	RASS	Arousal/agitation	10
Kim et al (2018)	Prospective cohort study	Hospital	224	0	69.3 (10.6)	68, 156	DSM-4 and CAM	RASS and NUDESC	Arousal/agitation, NUDESC sub-items	3
Lagarto et al (2020)	Prospective cohort study	Hospital (geriatrics)	227	82 (36.1%)	81.0 (7.9)	0, 269	CAM	RASS	Arousal/agitation	4
Liu et al (2018)	Randomised controlled trial	Hospital (acute)	100	Not reported	31.0 (SD not reported)	46, 54	CAM-ICU	SAS	Arousal/agitation	5
Maneewong et al (2017)	Prospective cohort study	Hospital (neuro-surgical)	25#	Not reported	38.9 (15.5)	4, 21	DSM-5	DRS-R98	DRS-R98 sub-items	4
Matsuda et al (2016)	Retrospective cohort study - chart review	Hospital (acute)	15	0	64.1 (9.5)	2, 13	DSM-4-TR	DRS-R98	DRS-R98 sub-items	2

Mercadante et al (2019)	Prospective cohort observational study	Acute supportive/palliative care unit	75††	Not reported	68.8 (11.1)	24, 62¶	MDAS	ESAS	ESAS sub-items	2
Tang et al (2018)	Prospective case-control study	Hospital (ICU)	21	0	61.0 (median) (mean and SD not reported)	8, 13	CAM-ICU	DelApp-ICU	Attention	4
Yang et al (2012)	Randomised controlled trial	Hospital (acute)	36	Not reported	69.8 (SD not reported)	13, 23	DSM-4	DRS-R98	DRS-R98 sub-items	6
FUNCTIONAL ABILITY										
Boettger et al (2011) §	Prospective cohort study	Cancer centre, psychiatry service	21	5 (23.8%)	69.6 (11.9)	11, 10	MDAS	KPS	Functional ability	3
Boettger et al (2011) §	Prospective cohort study	Cancer centre, psychiatry service	42	Not reported	69.6 (11.9)	Not reported	DSM-4-TR and MDAS	KPS	Functional ability	3
Boettger et al (2011) §	Secondary analysis of prospective cohort study	Cancer centre, psychiatry service	111	22 (19.8%)	23-89 (mean and SD not reported)	46, 65	DSM-4-TR and MDAS	KPS	Functional ability	3
Boettger et al (2014) §	Prospective cohort study	Cancer centre, psychiatry service	111	22 (19.8%)	23-89 (mean and SD not reported)	46, 65	DSM-4-TR and MDAS	KPS	Functional ability	3
Boettger et al (2014) §	Prospective cohort study	Cancer centre, psychiatry service	111	22 (19.8%)	23-89 (mean and SD not reported)	46, 65	DSM-4-TR and MDAS	KPS	Functional ability	3
Boettger et al (2015) §	Secondary analysis of prospective cohort study	Cancer centre, psychiatry service	111	22 (19.8%)	23-89 (mean and SD not reported)	46, 65	DSM-4-TR and MDAS	KPS	Functional ability	3

Flaherty et al (2010)	Retrospective observational study - chart review	Hospital (geriatrics)	44	Not reported	85.3 (5.7)	30, 14	CAM	Assessment of ADL (not specific tool)	Functional ability (ADLs)	2

† Lam et al, 2014 and Chong et al, 2015 reporting same study ‡ Meagher et al 2012 and Leonard et al 2013 reporting same study § Boettger et al 2011, 2011, 2011, 2014 reporting same study ¶ data for total sample (delirious and non-delirious participants) #18/25 included in results †† 86 delirious participants, only 75 followed for repeat assessment.

DRS-R98 = Delirium Rating Scale; CTD = Cognitive Test for Delirium; EDTB-ICU = Edinburgh Delirium Test Box; MMSE = Mini Mental State Examination; MoCA = Montreal Cognitive Assessment; SPMSQ = Short Portable Mental Status Questionnaire; KPS = Karnofsky Scale of Performance Status; TCT = Trunk Control Test; PANSS-P = Positive and Negative Syndrome Scale; NUDESC = Nursing Delirium Screening Scale; SAS = Sedation Agitation Scale; ESAS = The Edmonton Symptom Assessment Scale; CGBRS = Crichton Geriatric Behavioural Rating Scale.

Location in manuscript: following Results section 'assessment tools'.

Table 2: Assessment tools used in included studies

Repeat assessment tool	Specific domains	Range of scores	No. studies that used assessment	Cut-off for delirium †
Delirium-specific tools				
Delirium Rating Scale (DRS-R98) (individual items)	Sleep-wake cycle disturbance, perceptual disturbance and hallucinations, delusions, lability of affect, language, thought process abnormalities, motor agitation, motor retardation, orientation, attention, short-term memory, long-term memory, visuospatial ability	0-40	9	18
Cognitive Test for Delirium (CTD)	General cognition	0-30	3	18
Edinburgh Delirium Test Box (EDTB-ICU)	Attention	0-11	1	N/A
DelApp-ICU	Attention	0-12	1	N/A
Tools not specific for delirium				
Mini Mental State Examination (MMSE)	General cognition	0-30	15	N/A
Montreal Cognitive Assessment (MoCA)	General cognition	0-30	1	N/A
Clock Drawing Test	General cognition	0-1	1	N/A
Digit Span Task	General cognition	0-7	2	N/A
Cognitive Estimation Task	General cognition	15-60	1	N/A
Verbal Abstraction Task	General cognition	0-40	1	N/A

Short Portable Mental Status Questionnaire (SPMSQ)	General cognition	0-10	1	N/A
Karnofsky Scale of Performance Status	Functional ability	0-100	5	N/A
Trunk Control Test	Functional ability	0-100	1	N/A
Tinetti Scale	Functional ability	0-28	1	N/A
Modified Barthel Index	Functional ability	0-100	1	N/A
Non-specific assessment of ADL	Functional ability (ADLs)	N/A	1	N/A
Positive and Negative Syndrome Scale (PANSS-P)	Psychotic symptoms (disorganised thoughts, hallucinations, delusions, grandiosity, hostility, excitement and suspiciousness)	7-49	2	N/A
Nursing Delirium Screening Scale (NUDESC) (sub-items)	Disorientation, communication, hallucinations, motor ability	0-10	1	N/A
Sedation agitation scale (SAS)	Arousal/agitation	1-7	1	N/A
The Edmonton Symptom Assessment Scale (ESAS) (sub-items)	Depression, anxiety, drowsiness and non-relevant domains	0-10	1	N/A
Crichton Geriatric Behavioural Rating Scale (CGBRS)	Abnormal behaviour/ADLs	1-5	1	N/A

† Cut-off scores for delirium based on original validation studies, may vary across included studies.

Location in manuscript: following Results section 'assessment tools'.

Table 3: Table of findings

Study ID	Mean scores (baseline - last assessment)	Key findings	Significance	Statistics (change over time)	Conclusion of study	Longitudinal change/group differences
Bellelli et al (2011)	TCT: 31.40 - 65.95 Tinetti: 5.45 - 12.0 CDT: 0.8 - 1.95 DST: 2.7 - 1.10 CET: 1.30 - 2.00 VA: 0.65 - 1.65	Significant improvement in motor performance over time	All domains (apart from CDT (DSD)): $p < .05$ CDT (DSD): $p = .60$	Paired t-test	Patients with delirium and DSD exhibited a pattern of fluctuating motor performance and general cognition related to the onset and end of delirium.	Longitudinal change and group differences
Boettger et al (2011)	KPS: 28.10 - 41.00	Significant improvement in KPS score over time	KPS: $x^2 = 20.11$, $p < .001$	Friedman test	Delirious patients treated with aripiprazole showed improved functional ability over time.	Longitudinal change
Boettger et al (2011)	KPS: 25.25 - 36.45	Significant improvement in KPS scores over time for both groups. No significant difference in KPS scores between groups	KPS (aripiprazole group): $x^2 = 20.11$, $p < .001$ KPS (haloperidol group): $x^2 = 20.83$, $p < .001$	Friedman test	No difference in efficacy between aripiprazole and haloperidol in the improvement of delirium symptoms.	Longitudinal change and group differences
Boettger et al (2011)	KPS: 23.75 - 31.65	Significant improvement in KPS scores over time for both groups. No significant difference in KPS scores between groups	KPS (delirium and dementia): $x^2 = 17.54$, $p < .001$ KPS (delirium) = $x^2 = 76.56$, $p < .001$	Friedman test	The response to antipsychotic treatment did not differ between delirium sub-types.	Longitudinal change and group differences
Boettger et al (2014)	KPS: 24.00 - 33.00	Significant improvement in KPS score over time	KPS (hypoactive): $x^2 = 36.36$, $p < .001$ KPS (hyperactive): $x^2 = 57.65$, $p < .001$	Friedman test	Patients showed an improvement in functional ability with antipsychotic treatment, though DSD patients showed lower response rates.	Longitudinal change and group differences
Boettger et al (2014)	KPS: 24.1 - 33.0	Significant improvement in KPS score over time	KPS: $F = 80.5$, $p < .001$	ANOVA	Delirium had an acute impact on the level of functioning which was reversible with appropriate management of delirium.	Longitudinal change and group differences

Boettger et al (2015)	KPS: 23.67 - 31.2	Significant improvement in KPS score over time, for those with persistent delirium, resolved delirium at T2 and resolved delirium at T3	KPS: $p < .015$	Friedman test	Advanced age, dementia, brain cancer, terminal illness, infection and delirium severity were associated with prolonged and refractory course of delirium and lower functional status at 1 week of antipsychotics.	Longitudinal change and group differences
Breitbart et al (1996)	MMSE: 13.18 - 14.59	Significant improvement in MMSE scores over time only for chlorpromazine group	MMSE (chlorpromazine): $p < .04$ MMSE (haloperidol and lorazepam): $p > .05$	ANOVA	Early intervention with neuroleptic agents (chlorpromazine) in low doses may be useful in managing delirium in AIDS patients, but lorazepam led to increased cognitive impairment.	Longitudinal change
Chong et al (2015)	Graph only	Significant improvement in CMMSE scores over time	CMMSE: $p < .001$ †	Wilcoxon Signed Ranks test	There was a slower cognitive symptom recovery in patients with delirium superimposed on dementia, suggesting a cognitive reserve play role in delirium development and recovery.	Longitudinal change
Cole et al (1994)	SPMSQ: 8.20 - 8.00 CGBRS: 32.00 - 29.30	Significant improvement in SPMSQ scores, but not CGBRS, between groups over time.	SPMSQ: $F = 2.47$, $p < .05$ CGBRS: $F = 2.30$, $p = .06$	Multivariate analysis of variance	The beneficial effects of a geriatric service for detection and intervention in cases of delirium were small.	Group differences
Deschodt et al (2012)	MMSE: 4.32 - 4.96	Significant difference in MMSE scores between delirious group and non-delirious group.	MMSE: $p = .04$	Marginal linear model	The IGCT intervention reduced the incidence of adverse outcomes and geriatric consultation had no effect on the severity or duration of delirium episode, although more control than intervention participants had cognitive decline.	Group differences
Flaherty et al (2010)	ADL: 4.10 - 6.10	Significant improvement in ADLs over time (multivariate analyses controlled for covariates such as age etc. for ADL interaction effect)	ADL: $p < .001$	Mixed model analysis of variance	ACE unit with a delirium room may improve ADL function from admission to discharge, as shown by a significant interaction effect for delirious patients.	Longitudinal change and group differences

Gagnon et al (2005)	MMSE: 20.90 - 27.80	Significant improvement in MMSE scores over time	MMSE: $p < .05$	Matched, paired Wilcoxon signed rank test	Patients showed improvement in alertness, psychomotor retardation and slurred speech and general cognition with methylphenidate treatment.	Longitudinal change
Green et al (2017)	Graph only	Significant difference in EDTB-ICU scores between delirious and non-delirious patients at assessments 1-3	EDTB-ICU: $p < .001$	ROC analyses and linear effects model	EDTB-ICU performance was associated with delirium status and has the potential to track attentional deficits in delirious patients over time.	Group differences
Hui et al (2017)	Graph only	Significant difference in RASS scores between treatment group and control group	RASS: $p < .001$	Wilcoxon rank sum test	Treatment of lorazepam and haloperidol combined resulted in a significant reduction of agitation at 8 hours in delirious patients.	Group differences
Kim et al (2018)	RASS: 1.41 - 0.54 NuDESC: 4.07 - 2.98	Significant differences in scores over time in all four subtype groups (hyperactive sample largest - reported here)	RASS: $F = 9.66$, $p < .001$ NuDESC: $F = 5.11$, $p < .001$ ǂ	Paired t-test (in each subtype group)	Symptom fluctuation was a core feature of delirium and patterns of fluctuation within domains differed among subtypes.	Longitudinal change and group differences
Kishi et al (2012)	MMSE: 17.20 - 18.80 DRS item 1: 2.0 - 1.5 DRS item 2: 1.8 - 0.9 DRS item 3: 1.6 - 0.8 DRS item 4: 1.5 - 0.8 DRS item 5: 1.3 - 0.9 DRS item 6: 1.5 - 1.3 DRS item 7: 1.2 - 0.7 DRS item 8: 1.1 - 1.2 DRS item 9: 1.3 - 1.0 DRS item 10: 1.8 - 1.3 DRS item 11: 1.9 - 1.4 DRS item 12: 1.7 - 1.6 DRS item 13: 1.1 - 1.0	Significant improvement in DRS-R98 scores (except thought processes, motor retardation, LTM and visuospatial ability) and MMSE score over time	DRS: all sig ($p \leq .04$) except thought processes ($p = .08$), motor retardation ($p = .60$) and LTM ($p = .66$) and visuospatial (.74) MMSE: $p = .051$	Paired t-test and regression analysis	Risperidone is effective in treatment of delirium in advanced cancer patients: most DRS-R98 scores significantly improved and MMSE had a trend toward a significant improvement.	Longitudinal change

Lagarto et al (2020)	Graph only	No analyses of change in RASS scores over time.	No specific analyses of assessments <7 days	T-test and Mann-Whitney test	Patients with dementia particularly prone to manifest with acute changes in mental status, including delirium and moderate-severe sedation.	Group differences
Lam et al (2014)	C-MMSE: 5.72 - 9.16 MBI: 28.79 - 47.97	Significant improvement in MMSE score over time. Change in Modified Barthel Index score not significant.	MMSE: $p < .001$ MBI: $p = .28$	Pearson chi-square, Mann Whitney U test and independent-sample t-test	Patient with residual subsyndromal delirium had prolonged recovery trajectory of delirium.	Longitudinal change and group differences
Leonard et al (2013)	DRS item 1: 1.6 - 1.7 DRS item 2: 1.0 - 0.8 DRS item 3: 0.5 - 0.4 DRS item 4: 0.8 - 0.8 DRS item 5: 1.1 - 1.0 DRS item 6: 1.4 - 1.4 DRS item 7: 1.0 - 0.8 DRS item 8: 1.0 - 1.1 DRS item 9: 1.3 - 1.3 DRS item 10: 2.0 - 2.0 DRS item 11: 1.6 - 1.6 DRS item 12: 1.2 - 1.3 DRS item 13: 1.9 - 1.9 DRS item 14: 1.8 - 1.5 DRS item 15: 1.1 - 0.9 DRS item 16: 1.6 - 1.6*	No specific analyses of change from T1 to T2 (within 7-day timeframe). Interaction of individual symptoms with time did not have significant effects.	No specific analyses of assessments <7 days	Linear mixed effects model	Attention is disproportionately and consistently impaired throughout delirium episodes, all symptoms significantly contributed to DRS-R98 scores over time.	Longitudinal change
Lingehall et al (2017)	Graph only	No specific analyses of change from baseline to T2 or T3 (within 7-day timeframe).	No specific analyses of assessments <7 days	Independent samples t-test, Generalised estimating equations, univariate logistic regression analyses and multivariate models	Older patients with reduced pre-operative cognitive functions and those who develop delirium are at risk of dementia development during 5 years after cardiac surgery.	Group differences

Liu et al (2018)	Graph only	No analyses of change in SAS scores. SAS scores significantly higher in group D (DEX at loading dose for 10 mins then pumped for maintenance) than other groups.	SAS: $p < .05$	Independent samples t-test, ANOVA and chi-square (group differences)	DEX and sufentanil decrease incidence of post-operative delirium.	Group differences
Lou et al (2003)	MMSE: 22.13 - 20.88 Orientation time: 3.08 - 2.43 Orientation place: 4.30 - 3.95 Registration: 2.93 - 2.90 Language: 7.49 - 7.35 Recall: 2.28 - 2.43 Calculation: 3.60 - 3.52 Visual construction: 0.69 - 0.60	Significant improvement in MMSE scores over time	MMSE: $p < .01$	Repeated measures ANOVA	Participants with delirium had significantly lower MMSE scores and all sub-items than non-delirious participants in every stage.	Longitudinal change and group differences
Lowery et al (2008)	MMSE: 21.89 -22.63	No significant change in MMSE over time	MMSE: $F = 1.86$, $p = 0.178$	Repeated measures ANOVA	Attention and fluctuating cognition may offer excellent discriminative utility for delirium, future studies could benefit from the application of neuropsychological measures.	Longitudinal change and group differences
Maneeton et al (2013)	Graph only	No significant change in DRS-R98 cognitive sub-score over time	DRS-R98 cognitive sub-score: $p = .89$	Mixed model for repeated measurements (MMRM)	Low doses for both quetiapine and haloperidol are equally effective and safe for the management of behavioural disturbance in delirious patients, given together with environmental manipulation.	Longitudinal change and group differences

Maneewong et al (2017)	Graph only	Significant improvement in DRS-R98 sub-items (orientation, attention, visuospatial ability, lability of affect, motor agitation) over time	Orientation: $x^2 = 11.02$, $p = .01$ Attention: $x^2 = 10.16$, $p = .02$ Visuospatial ability: $x^2 = 14.91$, $p < .01$ Lability of affect: $x^2 = 13.71$, $p < .01$ Motor agitation: $x^2 = 12.73$, $p < .01$	Friedman test	Most symptoms of the DRS-R98 were prominent within the first day of TBI with delirium, symptoms that rapidly resolved included orientation, attention, visuospatial ability, lability of affect and motor agitation.	Longitudinal change and group differences
Matsuda et al (2016)	DRS item 1: 2.30 - 1.30 DRS item 2: 1.70 - 1.20 DRS item 3: 2.20 - 0.40 DRS item 4: 2.30 - 0.90 DRS item 5: 1.90 - 1.10 DRS item 6: 2.10 - 0.50 DRS item 7: 2.50 - 0.50 DRS item 8: 0.90 - 1.10 DRS item 9: 2.00 - 1.30 DRS item 10: 2.30 -1.30 DRS item 11: 1.80 -1.50 DRS item 12: 1.50 -1.10 DRS item 13: 1.90 - 1.10*	Significant improvement in DRS-R98 sub-items (except for motor retardation, short-term memory and long-term memory disturbance) over time	All sig ($p \leq .01$) except motor retardation ($p = 0.50$), STM ($p = 0.33$) and LTM ($p = 0.11$)	Paired t-test and regression analysis	Risperidone monotherapy is effective for treating delirium in patients with advanced cancer, in treating specific symptoms, except motor retardation, short-term memory disturbance, long-term memory disturbance.	Longitudinal change and group differences

Meagher et al (2012)	Graph only	Significant predictor variables for persistent delirium were orientation, memory, delusions, motor agitation, inattention and thought process abnormalities.	CTD: p values ≤ .030 DRS-R98 items: p values ≤ .038	Generalised estimating equations (GEE) method for longitudinal data for patterns in items from DSM/CTD.	Disturbance of attention and disturbed thinking were dominant elements throughout the course of delirium, particularly of persistent full syndromal delirium.	Group differences
Mercada-nte et al (2019)	Depression: 3.80 - 2.60 (not recovered) and 3.80 - 1.30 (recovered) Drowsiness: 3.90 - 3.40 (not recovered) and 4.00 - 2.60 (recovered) Anxiety: 3.80 - 1.70 (recovered) and 3.20 - 2.50 (not recovered)	Significant reduction in depression over time (non-recovered patients) and significant reduction in depression and anxiety over time (recovered patients). No significant change in drowsiness over time.	Depression: p = .001 (recovered); p = .017 (not recovered) Drowsiness: p = .071 (recovered); p = .343 (not recovered) Anxiety: p = .001 (recovered); p = .112 (not recovered)	Paired samples t-test	Patients admitted to APCSU with delirium reported higher distress in pain and depression. Patients who subsequently developed delirium after a week presented an even higher expression of multiple symptoms.	Longitudinal change and group differences
Milisen et al (2001)	MMSE: 14.30 - 13.60 (intervention group) and 7.00 - 8.90 (non-intervention group)	Significant improvement in memory and cognitive functioning over time; no significant difference between groups	No specific analyses of assessments <7 days	Linear mixed model	An integrated geriatric care model for delirium regarding severity and duration of the psychiatric symptoms benefitted older hip fracture patients, with a significant improvement in general cognition in all groups.	Longitudinal change and group differences
Mittal et al (2004)	CTD: 7.10 - 16.90 KSPS: 32.00 - 45.50	Significant improvement in CTD scores and KSPS scores over time	CTD: p = .0078 KSPS: p = .044	Paired t-test, F-test, chi-square	Risperidone is an effective and safe alternative to conventional antipsychotics in the treatment of delirium, particularly in improving general cognition and functional ability.	Longitudinal change
Parellada et al (2004)	MMSE: 13.13 - 26.38 PANSS-P: 21.50 - 10.14	Significant improvement in MMSE scores and PANSS-P scores over time	MMSE and PANSS-P: p < .001	Wilcoxon test and Friedman test	Symptoms of delirium in medically hospitalised patients may be treated in medical settings efficaciously and safely using risperidone at low doses.	Longitudinal change

Pintor et al (2009)	MMSE: 17.03 - 24.06 PANSS-P: 4.57 - 9.35	Significant improvement in MMSE scores and PANSS-P scores over time	MMSE: F= 96.56, p < 0.001 PANSS-P: F = 144.83, p < 0.001	MANOVA	Low doses of amisulpride may improve delirium symptoms in medical and surgical patients.	Longitudinal change
Saczynski et al (2012)	MMSE: 18.10 - 23.*30*	Significant improvement in MMSE scores from delirium diagnosis to repeat assessment	MMSE: p < .001 *§*	Hierarchical linear regression model	Postoperative development of delirium is a risk factor for decline in cognitive function and a prolonged period of impairment after surgery.	Longitudinal change
Tahir et al (2010)	DRS-R98 cog: 8.88 - 6.11 MMSE: 11.83 - 16.55	Significant difference in DRS-R98 cognitive score but not in MMSE score between treatment and control group	DRS-R98 cog: p < .001 MMSE: p = .197	Non-linear mixed effects model (group differences)	Quetiapine appeared to be well tolerated treatment for delirium with no evidence of significant adverse effects.	Group differences
Tang et al (2018)	DA-ICU: 0.50 - 0.50 (medians)	Significant difference in DelApp-ICU score between delirious and non-delirious group	DA-ICU: p < .001	Mann-Whitney U test, Chi-square test (group differences)	The DelApp-ICU shows promise as an objective tool to assist detection of delirium, the diagnosis of delirium was associated with a decrease in DelApp-ICU score.	Group differences
William et al (2017)	MoCA: 2.85 - 3.12	Higher MoCA scores were independent predictors for delirium recovery	MoCA: p = .026	Generalised estimating equations models (GEE)	BDNF levels could be a marker of delirium recovery, suggesting recovery may be predicted based on biological factors as well as higher MOCA scores.	Group differences
Yang et al (2012)	Graph only	Mean DRS-R98 scores decreased (i.e., improved) and significant time-by-treatment group interaction for DRS scores	DRS-R98: F = 2.87, p = .025	Repeated measures ANCOVA	Adjuvant bright light therapy with risperidone might be useful for improving delirium and sleep-wake cycle disturbance.	Group differences

Yoon et al (2013)	MMSE: 15.10 - 22.80 DRS-R98 cog: 8.14 - 3.94	Significant improvement in MMSE score and DRS-R98 cognitive sub-score over time	MMSE and DRS-R98 cog: $p < .001$ (within-groups); $p > .565$ (between-groups)	Linear mixed model	Risperidone, olanzapine, quetiapine and low doses of haloperidol were equally effective and safe in the treatment of delirium.	Longitudinal change and group differences

† Adjusted (for variables that differed at baseline) ‡ Adjusted (for RM ANOVA to assess differences in RASS and NuDESC scores based on interaction effects for time x subtype group) § Adjusted (for age, educational level, sex, race or ethnic group, score on Charlson comorbidity index, presence or absence of a history of stroke or transient ischemic attack)

DRS-R98 = Delirium Rating Scale; CTD = Cognitive Test for Delirium; EDTB-ICU = Edinburgh Delirium Test Box; MMSE = Mini Mental State Examination; MoCA = Montreal Cognitive Assessment; SPMSQ = Short Portable Mental Status Questionnaire; KPS = Karnofsky Scale of Performance Status; TCT = Trunk Control Test; PANSS-P = Positive and Negative Syndrome Scale; NUDESC = Nursing Delirium Screening Scale; SAS = Sedation Agitation Scale; ESAS = The Edmonton Symptom Assessment Scale; CGBRS = Crichton Geriatric Behavioural Rating Scale.

Location in manuscript: following Results section 'assessment tools'.

Figure 1 legend: PRISMA flow diagram of study selection

Figure in file ‘Figure1.pdf’

Appendix 1, 2, 3 and 4 in file ‘Appendices.docx’.