Archives of Gerontology and Geriatrics 83 (2019) 292–302

Contents lists available at ScienceDirect

Archives of Gerontology and Geriatrics

journal homepage: www.elsevier.com/locate/archger

Review

Delirium risk in non-surgical patients: systematic review of predictive tools

Dylan Flaws[a], Gerard Byrne[a,b], Alison M Mudge[b,c,*]

[a] Royal Brisbane and Women's Hospital Mental Health Services, Brisbane, Australia
[b] Faculty of Medicine, University of Queensland, Brisbane, Australia
[c] Royal Brisbane and Women's Hospital, Department of Internal Medicine and Aged Care, Brisbane, Australia

ARTICLE INFO

Keywords:
delirium
clinical prediction rule
systematic review

ABSTRACT

Objective: Delirium is a common, serious condition associated with poor hospital outcomes. Guidelines recommend screening for delirium risk to target diagnostic and/or prevention strategies. This study critically reviews multicomponent delirium risk prediction tools in adult non-surgical inpatients.

Study design: Systematic review of studies incorporating at least two clinical factors in a multicomponent tool predicting risk of delirium during hospital admission. Derivation and validation studies were included. Study design, risk factors and tool performance were extracted and tabulated, and study quality was assessed by CHARMS criteria.

Data sources: PubMed, Embase, PsycINFO, and Cumulative Index to Nursing Health Literature (CINAHL) to 11th March 2018.

Data synthesis: 22 derivation studies enrolling 38,874 participants (9 with a validation component) and 4 additional validation studies were identified, from a range of ward types. All studies had at least moderate risk of bias. Older age and cognitive, functional and sensory impairment were important predisposing factors. Precipitating risk factors included infection, illness severity, renal and electrolyte disturbances. Tools mostly did not differentiate between predisposing and precipitating risk factors mathematically or conceptually Most tools showed fair to good discrimination, and identified more than half of older inpatients at risk.

Conclusions: Several validated delirium risk prediction tools can identify patients at increased risk of delirium, but do not provide clear advice for clinical application. Most recommended cut-points are sensitive but have low specificity. Implementation studies demonstrating how risk screening can better direct clinical interventions in specific clinical settings are needed to define the potential value of these tools.

1. Introduction

Delirium is a syndrome characterised by the acute onset of impaired attention and awareness and changes in cognition due to a precipitating condition, such as medication or illness (Inouye, 2006). It is common among adult inpatients and increases in older age, with up to 37% of medical inpatients over the age of 65 experiencing delirium during their admission (Cole et al., 2002; Eeles et al., 2010; Holden, Jayathissa, & Young, 2008; Inouye, 2006; Iseli, Brand, Telford, & LoGiudice, 2007; Lundström et al., 2005). Delirium may be present at the time of hospital presentation (prevalent) or develop during the hospital admission (incident). In either case, delirium is associated with an increased risk of poor hospital outcomes including falls, prolonged length of stay, functional decline, nursing home placement, readmission and death (Inouye, Rushing, Foreman, Palmer, & Pompei, 1998; Marcantonio, 2017; O'Keeffe & Lavan, 1997; Witlox et al., 2010).

Delirium can present in a similar manner to dementia, depression, and acute psychiatric syndromes, and these conditions can co-exist with delirium (Marcantonio, 2017). Early, accurate diagnosis is important to inform appropriate investigation and management. Even more importantly, multicomponent non-pharmacological prevention programs can reduce delirium incidence in patients at risk by up to one third (Marcantonio, 2017). Early identification of patients who might benefit from a diagnostic screen and/or preventive interventions may assist clinicians and researchers.

Risk factors for delirium are well established and include predisposing factors such as older age, pre-existing cognitive impairment, and sensory impairment as well as precipitating factors such as medications and severe medical illness (Elie, Cole, Primeau, & Bellavance, 1998; Inouye, 1998). Delirium is typically multifactorial, with multiple risk

* Corresponding author at: Internal Medicine Research Unit, Level 6 Block 7, Royal Brisbane and Women's Hospital, Butterfield St, Herston, Queensland 4029, Australia.
E-mail address: Alison.Mudge@health.qld.gov.au (A.M. Mudge).

https://doi.org/10.1016/j.archger.2019.05.013
Received 30 September 2018; Received in revised form 9 April 2019; Accepted 14 May 2019
Available online 18 May 2019
0167-4943/

factors producing a cumulative effect (Marcantonio, 2017) [10]. Multicomponent predictive tools which stratify risk based on predisposing and/or precipitating risk factors could identify patients at risk of prevalent or incident delirium (Siddiqi, 2016) to direct assessment and preventive efforts.

Clinical practice guidelines identify the importance of risk screening based on known risk factors but do not recommend specific tools or cut-points (Tropea, Slee, Brand, Gray, & Snell, 2008; Young, Murthy, Westby, Akunne, & O'Mahony, 2010). Two previous reviews (Lindroth et al., 2018; Newman, O''Dwyer, & Rosenthal, 2015) have shown that several multicomponent tools have good discrimination in various hospital inpatient populations, but they do not provide detailed analysis of classification characteristics of different tools (such as sensitivity, specificity, and positive and negative predictive values). It is important to consider whether a delirium prediction tool is highly sensitive (separating a smaller low risk group from a larger intermediate-high risk group) or highly specific (separating a smaller high risk group from a larger low-intermediate risk group), as the roles and resource implications of such tools may differ.

The objective of this systematic review was to critically review all studies which have derived multicomponent prognostic prediction tools for delirium in adult non-surgical inpatients, and any external validations or studies that aimed to modify/update/simplify such tools. In addition to critically appraising quality of the studies, describing the risk factors considered during derivation, and comparing the performance of the predictive tools, the review has a particular focus on classification measures and their potential implications for practice.

2. Materials and methods

This systematic review was prospectively registered on PROSPERO (registration number CRD42016053750).

2.1. Search strategy and selection criteria

Predictive tools were defined as any model or algorithm that incorporated at least two clinical factors measured at or during hospital admission in order to assign an estimated risk of developing delirium during the acute hospital stay. We included studies which reported prevalent, incident or unspecified onset delirium, recognising the challenge of clearly distinguishing early onset incident delirium in clinical settings. Observational studies (including cohort and case-control studies) were eligible if they derived a predictive tool in an adult inpatient setting, or aimed to validate or modify a previously derived tool. Studies were excluded if they were not available in English, did not report a predictive tool, examined delirium as a risk factor for other outcomes, or reported a diagnostic rather than predictive tool. Post-operative and intensive care settings were excluded because there are likely to be risk factors specific to these settings. We also excluded studies confined to specific clinical syndromes (delirium tremens, post-stroke delirium).

PubMed, Embase, PsycINFO, and Cumulative Index to Nursing Health Literature (CINAHL) were searched from earliest available records to 11th March 2018 for titles or abstracts containing either Delirium or Acute Confusion* as well as Predict*, Risk, Tool, Estimator, Algorithm, or Decision Aid. Duplicate entries were removed, then titles and abstracts of remaining records were screened independently by two reviewers. Records meeting criteria or with insufficient information were retrieved for full text review, with consultation with the third author in cases of uncertainty. Citations of all included studies were reviewed for any additional studies.

2.2. Data extraction and synthesis

Data were collected independently from published studies and supplements by two authors. Data included year of publication and authors, study design, sample size and participants, definition and measure of delirium, number of delirium cases, definition and measurement of predictor variables, how missing data were handled, derivation process for the predictive model, any post-derivation modifications, prediction tool structure and performance (calibration, discrimination and classification), validation methods, and results reporting. Where specified, prevalent and incident delirium were reported separately.

Quality was assessed for each study by two independent reviewers using CHARMS criteria (Moons et al., 2014). This approach includes nine domains (source, outcome, predictors, sample size, missing data, model development, tool performance and tool evaluation), with model development excluded for papers reporting only external validation of a previous tool. For each domain, studies were assessed as being either at low, medium, or high risk of introducing bias using collaboratively developed criteria.

Descriptive summaries including study design, population sample, tool structure and performance were tabulated for all derivation and validation studies. Risk factors included in each derivation study were tabulated, noting whether each risk factor was incorporated into the final tool. Risk factors were grouped into categories for descriptive purposes.

The proportion of participants with delirium was described for each study, and for each risk category within the final tools as reported by the authors. Discrimination was described using performance statistics reported by the authors. Classification was summarised using sensitivity, specificity, positive predictive value and negative predictive value at the recommended cut-point(s) of each tool, either as reported by the authors or calculated from data provided within the study report.

3. Results

3.1. Study Characteristics

We identified 22 studies which developed new tools (Bannink, van Veluw, van Zuijlen, Enting, & van der Rijt, 2004; Carrasco, Villarroel, Andrade, Calderón, & González, 2014; de Wit et al., 2016; Douglas et al., 2013; Francis, Martin, & Kapoor, 1990; Han et al., 2009; Hare, Arendts, Wynaden, & Leslie, 2014; Inouye & Charpentier, 1996; Inouye, Viscoli, Horwitz, Hurst, & Tinetti, 1993; Isfandiaty, Harimurti, Setiati, & Roosheroe, 2012; Kennedy et al., 2014; Kobayashi, Takahashi, Arioka, Koga, & Fukui, 2013; Levkoff, Safran, Cleary, Gallop, & Phillips, 1988; J.A. Martinez et al., 2012; Neefjes et al., 2017; O'Keeffe & Lavan, 1996; Patten, Williams, Petcu, & Oldfield, 2001; Pendlebury, Lovett, Smith, Wharton, & Rothwell, 2017; Pompei et al., 1994; Rudolph, Doherty, Kelly, Driver, & Archambault, 2016; Sri-On, Tirrell, Vanichkulbodee, Niruntarai, & Liu, 2016; Tadiri, 2016) (including nine with external validation) and four additional validation studies (Brown et al., 2017; Mestres Gonzalvo et al., 2017; Pendlebury et al., 2016; Rudolph et al., 2011). Study selection is shown in Fig. 1. Study characteristics for derivation and validation studies are summarised in Tables 1 and 2. In the 22 derivation studies, 14 studies were prospective cohorts and eight were retrospective, including five cohort studies and three case-control (Table 1). One study (Rudolph et al., 2016) included both a retrospective and prospective cohort; the prospective study was considered validation. Sample size ranged from 100 to 27,625, and included participants from geriatric, medical, emergency medicine, psychiatric, oncology and palliative care wards. Many studies only included older adults, although age cut-offs were inconsistent. Most of the studies in geriatric and older medical populations focussed on incident delirium, while the emergency department studies reported prevalent delirium; most other studies did not specify time of onset. Delirium was identified either by the Confusion Assessment Method (CAM) (Carrasco et al., 2014; Douglas et al., 2013; Han et al., 2009; Hare et al., 2014; Inouye & Charpentier, 1996; Inouye et al., 1993; Kennedy et al., 2014; J.A. Martinez et al., 2012; Pompei et al., 1994; Sri-On et al., 2016) or

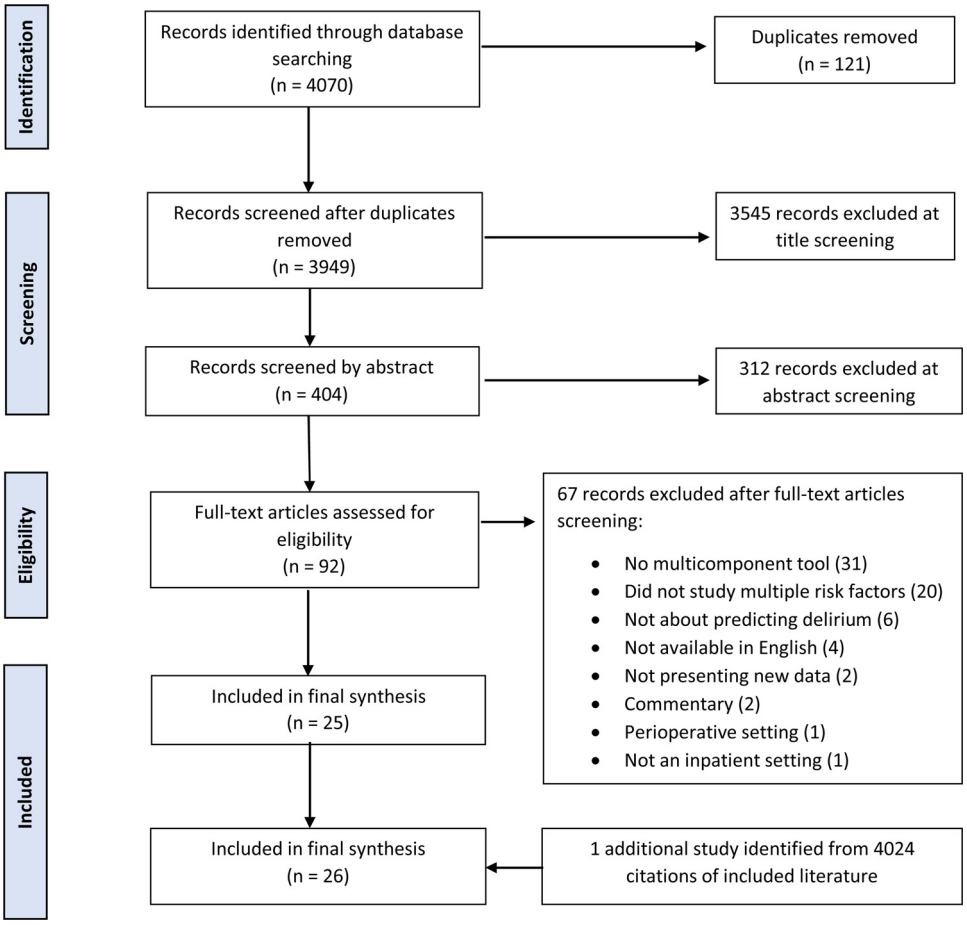

Fig. 1. Article Selection Process.

diagnostic assessment using Diagnostic and Statistical Manual (DSM) criteria (Francis et al., 1990; O'Keeffe & Lavan, 1996; Pendlebury et al., 2017) in prospective studies; retrospective studies relied on clinician diagnosis (Isfandiaty et al., 2012; Levkoff et al., 1988; Neefjes et al., 2017), coding in clinical or electronic records (de Wit et al., 2016; Rudolph et al., 2016), or application of indicators based on DSM criteria (Bannink et al., 2004; Kobayashi et al., 2013; Patten et al., 2001; Tadiri, 2016).

3.2. Risk of Bias Assessment

Risk of bias is summarised in supplementary Table S1. All studies had at least moderate risk of bias in more than one domain, and most had one or more area with high risk of bias. Retrospective studies were at high risk of outcome ascertainment bias, because delirium is often poorly documented (Inouye, 2006). Risk factors and outcomes were often collected by the same data collectors, selection of candidate risk factors was sometimes poorly justified, and variables were often dichotomised for analysis. Few studies reported missing data or how this was managed. Many studies did not report sample size calculations, and were at risk of overfitting because of relatively low number of events per variable (EPV, see Table 1). Model development was usually based on univariate analyses, and evaluation was often limited, without validation in an independent sample.

3.3. Risk factors

A large range of potential risk factors were included in the studies, using a variety of measures. Risk factors were collected from documentation (e.g. demographics, comorbidities, medications), self-report (e.g. ADL function), objective testing (e.g. cognitive testing, blood tests) and/or subjective clinical impressions (e.g. illness severity). There was considerable heterogeneity between studies in measurement methods. For example, cognitive impairment was measured directly in some prospective studies using a variety of standardised assessment tools such as the Blessed Dementia Rating Scale (Francis et al., 1990; O'Keeffe & Lavan, 1996), Mini Mental Status Examination (Inouye & Charpentier, 1996; Inouye et al., 1993; Pendlebury et al., 2017; Pompei et al., 1994; Sri-On et al., 2016), Pfeffer Functional Acitvities Questionnaire (Carrasco et al., 2014), Short Portable Mental Status Questionnaire (Sri-On et al., 2016), Abbreviated Mental Test (Pendlebury et al., 2017), mini-Cog (Douglas et al., 2013)or Informant Questionnaire on Cognitive Decline in the Elderly (Han et al., 2009), with several studies using multiple tools. Other prospective studies (Hare et al., 2014; Kennedy et al., 2014; J.A. Martinez et al., 2012; Tadiri, 2016) and retrospective studies (Bannink et al., 2004; Isfandiaty et al., 2012; Kobayashi et al., 2013; Neefjes et al., 2017; Patten et al., 2001; Rudolph et al., 2016) used evidence of documented cognitive impairment in medical records, although choices of included terms varied between studies.

To permit comparison, risk factors were grouped into domains of predisposing factors (patient characteristics which make them vulnerable to an insult) and precipitating factors (markers of illness or injury which may precipitate delirium). Supplementary Table 2 illustrates the variety of risk factors studied within each domain, and Table 3 summarises which domains were included in model development and which were included in the final model for each study. Cognitive impairment was included as a significant domain in almost all of the models which tested it. Functional status and sensory impairment were included in half the models which tested them, and older age was

Table 1
Model derivation, including sample characteristics, delirium rates, modelling approach, tool structure and discrimination. Where multiple delirium rates are reported in a study (any, prevalent and/or incident), the measure for which tool performance is reported is marked in bold. CAM confusion assessment method; DSM Diagnostic and Statistical Manual; AUC area under the curve (c statistic); EPV events per variable; UTI urinary tract infection; WBC white blod cell count; MDC major dia = gnotic categories; ECOG European Cooperative Oncology Group functional status

Study and Design	Sample	Delirium diagnosis (identified by)	Delirium cases identified, n (%) and type	Risk factors included in the tool (weighting)	Modelling approach	Reported risk levels	Observed delirium cases, by risk level	Discrimination and events per variable (EPV)
Older medical and geriatric inpatients								
Francis et al. (1990) Prospective cohort	n = 229 Age > 70 Medical	DSM III (researcher)	**50 (21.8) any** 36 (15.7) prevalent 14 (6.1) incident	Abnormal sodium level Illness severity Cognitive Impairment Fever or hypothermia Psychoactive drug use	Regression	0 Factors 1 Factor 2 Factors 3 Factors ≥4 Factors	3% 6% 32% 48% 100%	EPV 10
Inouye et al. (1993) Prospective cohort	n = 107 Age > 70 Medical	CAM (researcher)	**27 (25.2) incident**	Vision Impairment Severe Illness Cognitive impairment BUN/Cr ratio ≥18	Proportional hazards	0 Factors 1-2 Factors 3-4 Factors	9% 23% 83%	AUC 0.74 (0.63-0.85) EPV 7
Inouye and Charpentier (1996) Prospective cohort	n = 196 Age > 70 Medical	CAM (researcher)	**36 (17.9) incident**	Restraint Use Albumin < 30 g/l > 3 new medications Catheter Any iatrogenic event	Regression	0 Factors 1-2 Factors 3+ Factors	3% 20% 59%	Goodness of fit statistic 3.9 EPV 7
O'Keeffe and Lavan (1996) Prospective cohort	n = 100 "Elderly" Acute geriatrics	DSM III (researcher)	**28 (28.0) incident**	Dementia Severe Illness Urea > 10 mmol/L	Regression	0 Factors 1 Factor 2 Factors 3 Factors	8% 23% 60% 100%	AUC 0.79 EPV 9
Isfandiaty et al. (2012) Retrospective cohort	n = 457 Age > 60 Medical, acute geriatrics	Clinical team diagnosis	87 (18.8) incident	Sepsis (3.5) Infection (1) Decrease functional status (2) Cognitive impairment (3)	Proportional Hazards	Score < 2 Score 2-5 Score > 5	4% 33% 55%	AUC 0.82 (0.78-0.88) EPV 22
Douglas et al. (2013) Prospective cohort	n = 209 Age > 50 Medical	CAM (researcher)	25 (12.0) incident	Age > 80 Unable to spell DLROW Not oriented Illness severity	Regression	0 Factors 1 Factors 2 Factors 3 Factors 4 Factors	0% 5% 15% 26% 100%	AUC 0.81 (0.72-0.90) EPV 6
Carrasco et al. (2014) Prospective cohort	n = 374 Age > 65 Medical	CAM (researcher)	25 (6.7) incident	BUN/Cr ratio Barthel Index	Regression	Score < -240 (1370xBUN/creat - 4xBarthel index)		AUC 0.86 (0.82-0.91) EPV 13
Tadiri (2016) Prospective cohort	n = 520 Age > 65 Medical	Medical record using chart-based CAM	77 (14.8) incident	Age ≥ 80 (1) Dementia (3) History of stroke/ TIA (1) ED diagnosis stroke (2) ED diagnosis fall (1) O_2 Saturation < 90% (8) Sodium < 130mEq/L (2)	Regression	Score 1 Score 2 Score 3 Score 4+	11% 35% 38% 45%	AUC 0.73 EPV 11
Pendlebury et al. (2017) Prospective cohort	N = 308 Age > 65 Medical	DSM IV (treating physician and researcher) following routine CAM screening	**95 (30.8) any** 28 (9.1) incident 67 (21.8) prevalent	Cognitive impairment (2) Age ≥80 (2) Severe illness (1) Infection (1) Vision Impairment (1)	A priori based on systematic review	Score ≤1 Score 2-4 Score 5+	13% 30% 74%	AUC 0.78 (0.71-0.84) EPV 19
Older mixed hospital inpatients								
Levkoff et al. (1988) Retrospective case control	n = 1,285 Age > 60 Mixed	Medical record discharge coding	117 any Case: control ratio = 1:10	UTI Albumin < 3.4 mg/dL WBC	Recursive partitioning	Groups 1 UTI 2 Albumin < 3.4 3 WBC ≥11,000		EPV 29

(*continued on next page*)

Table 1 (*continued*)

Study and Design	Sample	Delirium diagnosis (identified by)	Delirium cases identified, n (%) and type	Risk factors included in the tool (weighting)	Modelling approach	Reported risk levels	Observed delirium cases, by risk level	Discrimination and events per variable (EPV)
	hospital inpatients		Identification rate 0.7%	≥11,000/mm^3 Proteinuria		4 Proteinuria All negative	19% 14% 10% 9% 4%	
Pompei et al. (1994) Prospective cohort	n = 432 Age > 65 Mixed hospital inpatients	DSM III after CAM (researcher)	**64 (14.8) any** 21 (4.9) prevalent 43 (10.0) incident	Sum of weighted variables Cognitive Impairment (2) 4+ MDCs (3) Depression (2) Alcoholism (3)	Regression	Score 0-3 Score 4-7 Score 8-10	3% 17% 33%	AUC 0.74 (0.70-0.78) EPV 13
de Wit et al. (2016) Retrospective case control	n = 1,291 Age > 60 Mixed hospital inpatients	Electronic medical record coding	225 incident case: control 4.7:1 Overall case identification rate not reported	Age (V1) Polypharmacy (V2) Anxiolytics (V3) Antidementia (V4) Antidepressants (V5) Antiparkinson's (V6) Antidiabetics (V7) Antipsychotics (V8) Analgesics (V9) Sleep medication (V10) CRP (V11) Urea (V12)	Regression	−9.118 +(0.082*V1) +(0.029*V2) +(0.248*V3) +(1.123*V4) +(0.286*V5) +(1.962*V6) +(0.330*V7) +(1.164*V8) +(0.371*V9) +(0.155*V10) +(0.002*V11) + (-0.001*V12)		AUC 0.78 (0.74-0.81) EPV 19
Rudolph et al. (2016) Retrospective cohort	n = 27,625 Age > 55 Mixed hospital inpatients	Electronic medical record coding	2,343 (8.5) prevalent	Cognitive Impairment (4) Age ≥65 (2) Age ≥80 (3) Infection (2) Fracture (4) Vision impairment (1) Severe Illness (2)	A priori based on systematic review	Score 0-2 Score 3-5 Score 6-8 Score ≥9	2% 3% 11% 38%	AUC 0.81 (0.80-0.82) EPV 335
Adult medical inpatients								
J.A. Martinez et al. (2012) Prospective cohort	n = 397 Age > 18 Medical	CAM (researcher)	52 (13.1) any	Age > 85 ADL dependence > 5 Antipsychotic or > 2 psychotropics	Regression	0 Factors 1 Factor 2+ Factors	4% 14% 43%	AUC 0.77 (0.73-0.82) EPV 17
Kobayashi et al. (2013) Retrospective cohort	n = 2,400 Adult Medical	Clinical team diagnosis	91 (3.8) any	Previous delirium Age Malignancy ADL	Chi square automatic interaction detector (CHAID) decision tree	Quite Low Low 1 Low 2 Moderate High Quite High	0% 1% 1% 3% 9% 40%	AUC 0.82 (0.77-0.86) EPV 23
Older emergency department patients								
Han et al. (2009) Prospective cohort	n = 303 Age > 65 Emergency dept	CAM-ICU (researcher)	25 (8.3) prevalent	Dementia ADL dependence Hearing Impairment	Regression	0 Factors 1 Factor 2 Factors 3 Factors	1% 8% 19% 46%	AUC 0.82 EPV 8
Hare et al. (2014) Prospective cohort	n = 320 Age > 65 Emergency dept	CAM (researcher)	23 (7.2) prevalent	History of dementia (2) History of depression (1) Arrhythmia (1)	Regression	Score 0 Score 1 Score 2 Score 3 Score 4	0% 2% 13% 22% 43%	AUC 0.86 EPV 8
Kennedy et al. (2014) Prospective cohort	n = 676 Age > 65 Emergency dept	CAM (researcher)	63 (9.3) prevalent	Age 70-79 (1) Age 80-89 (2) Age ≥90 (3) Dementia (3) History TIA or stroke (2) Respiratory rate > 20 (2) Suspected infection (2)	Regression	Score ≤2 Score 3-4 Score ≥5	4% 12% 29%	AUC 0.77 (0.71-0.83) EPV 8

(*continued on next page*)

Table 1 (*continued*)

Study and Design	Sample	Delirium diagnosis (identified by)	Delirium cases identified, n (%) and type	Risk factors included in the tool (weighting)	Modelling approach	Reported risk levels	Observed delirium cases, by risk level	Discrimination and events per variable (EPV)
				ED diagnosis of ICH (5)				
Sri-On et al. (2016) Prospective Cohort	n = 232 Age > 65 Emergency dept	CAM-ICU (researcher)	27 (11.6) prevalent	Auditory Impairment (13) Dementia (4.7) Metabolic derangement (6.5)	Regression			AUC 0.82 EPV 9
Adult inpatients in specific ward settings								
Patten et al. (2001) Retrospective case control	n = 425 Age not specified Psychiatric	Medical record discharge coding (ICD-9) (verified by researcher)	122 any 91 prevalent **31 incident** case: control = 4:1 (incident) and 2:1 (prevalent) Overall case identification rates not reported	Age > 65 (18.2) Cognitive Impairment (3.7) Manic episode 10.5) Lithium (2.2) Antipsychotics (9.0) Antichol/ antiparkinson medications (4.9)	Regression			EPV 20
Bannink et al. (2004) Retrospective cohort	n = 104 Age not specified Palliative care	DSM IV (clinician) after structured screening process	13 (12.5) any	Previous delirium (3) Cognitive impairment (3) Opioid change (3) Age > 70 (1) Alcohol excess (1) Recreational drugs (1) Fever > 38.5^{0}C (1) Metabolic problems (1) Hearing/vision problem (1) Anaesthesia (1)	A priori based on literature and clinical experience	Score < 3 Score ≥3	2% 24%	EPV 1
Neefjes et al. (2017) Retrospective Cohort	N = 620 Age not specified Oncology	Delirium Observation Screening Scale (nursing staff) or documented in medical record.	**98 any** 52 from original cohort 46 added to improve statistical power	Unscheduled admission Metabolic Imbalance ECOG Performance Palliative treatment	Recursive Partitioning	Admission 1. Unscheduled (Y/N) 2. Metabolic Imbalance (Y/N) 3. ECOG (0-2/3-4) 4. Treatment (Palliative/ Curative)	9% 16%/1% 10%/33% 54%/22% 18%/60%	AUC 0.81 EPV 25

included in nearly half studies where age was not included in the final model were confined to older patients. Importantly, all except two tools (de Wit et al., 2016; Levkoff et al., 1988) included functional or cognitive impairment, and the presence of either factor would classify the patient at intermediate or high risk. Findings among precipitating factors were less consistent. Infection, illness severity, and renal and electrolyte disturbances were tested most commonly and were included in 37-50% of resulting tools; low albumin, iatrogenic factors and new medications were seldom tested, but included in nearly half of studies testing them. New medications were only addressed in two studies but included in the final model on both occasions. Pain was only tested as a risk factor in one study.

3.4. Predictive tool development and performance

Table 1 summarises 22 studies (including 38,874 participants) developing a risk prediction tool. These tools contained between 2 and 10 risk factors, most commonly selected using logistic regression although other approaches included decision tree methods and time to event methods (Table 1). Three studies pre-specified the risk factors based on systematic review and/or clinical opinion without developing a statistical model (Bannink et al., 2004; Pendlebury et al., 2017; Rudolph et al., 2016). Earlier tools assigned scores based on the number of risk factors, but in later studies authors more commonly assigned weights derived from odds ratios to create a weighted risk score. Pendlebury did not find any significant improvement in tool discrimination when directly comparing weighted and unweighted approaches with the same risk factors (Pendlebury et al., 2017). Most studies created risk subgroups and reported observed delirium incidence or prevalence within these groups, but none provided calibration plots or slopes. Discrimination was most commonly described using area under the receiving operator curve (AUROC). Most models had fair to good discrimination, reporting AUROC between 0.69 and 0.86.

3.5. Validation studies

Validation studies of the predictive tools (14 studies enrolling 3,235 participants) are summarised in Table 2. Two derivation studies used an internal split sample and four used resampling methods within the derivation cohort for internal validation. External validation was most

Table 2
Model validation. The table shows the type of validation for each study, and details of external validation sample and predictive tool discrimination. AUC area under the curve (c statistic); CAM Confusion Assessment Method; DSM Diagnostic and Statistical Manual; DOSS delirium observation screening scale.

Derivation study	Validation study	Validation type	Sample	Delirium measure	Delirium rate in validation, %	Discrimination AUC (95% confidence interval)
Older medical and geriatric inpatients						
Francis et al., 1990		none				
Inouye et al., 1993	Inouye et al., 1993	External (temporal)	N = 174 Age > 70 Medical	CAM (researcher)	16.7 incident	0.66 (0.55-0.77)
	Rudolph et al., 2011	External (temporal, geographic)	N = 100 Age > 65 Medical	DSM-IV (researcher)	23.0 any	0.65 (0.54-0.76)
	Pendlebury et al., 2016	External (temporal, geographic)	N = 308 Age > 65 Medical	DSM IV (treating physician and researcher) following routine CAM screening	30.8 any 21.8 prevalent 9.1 incident	0.73 (0.66-0.80) any 0.70 (0.62-0.72) prevalent 0.73 (0.62-0.84) incident
Inouye and Charpentier, 1996	Inouye and Charpentier, 1996	External (temporal)	N = 312 Age > 70 Medical	CAM (researcher)	15.0 incident	Goodness of fit statistic 4.1
O'Keefe, 1996		Internal (split sample)				
Isfandiaty, 2012	Pendlebury et al., 2016	External (temporal, geographic)	N = 308 Age > 65 Medical	DSM IV (treating physician and researcher) following routine CAM screening	30.8 any 21.8 prevalent 9.1 incident	0.77 (0.70-0.83) any 0.69 (0.61-0.77) prevalent 0.83 (0.74-0.90) incident
Douglas, 2013	Douglas et al., 2013	External (temporal, geographic)	N = 165 Age > 50 Medical/surgical	CAM (researcher)	8.5 incident	0.69 (0.54-0.83)
	Pendlebury et al., 2016	External (temporal, geographic)	N = 308 Age > 65 Medical	DSM IV (treating physician and researcher) following routine CAM screening	30.8 any 21.8 prevalent 9.1 incident	0.74 (0.67-0.80) any 0.68 (0.60-0.76) prevalent 0.73 (0.63-0.84) incident
	Brown et al., 2017	External (temporal, geographic)	n = 347 Age > 50 Neurology	Medical record using chart-based CAM	19.9 any 14.1 prevalent 5.8 incident	0.83 (0.77-0.89) any 0.73 (0.60-0.85) incident
Carrasco et al., 2014	Carrasco et al., 2014	External (temporal)	N = 104 Age > 65 Medical	CAM (researcher)	11.5 incident	0.78 (0.66-0.90)
Tadiri, 2016		Internal (boot-strapping)				0.74
Pendlebury et al., 2017		none				
Older mixed hospital inpatients						
Levkoff et al., 1988	Levkoff et al., 1988	External (temporal)	N = 471 Age > 60 Medical, surgical, psychiatric	Electronic discharge diagnosis coding	Case control 43 cases, 428 randomly selected controls	Similar classification to validation cohort
Pompei et al., 1994	Pompei et al., 1994	External (temporal, geographic)	N = 323 Age > 70 Medical/surgical	CAM (researcher)	26.3 any 14.8 prevalent 11.8 incident	0.64 (0.59-0.69) any
deWit 2016		Internal (boot-strapping)				0.76
	Mestres Gonzalvo et al., 2017	External (temporal, geographic)	N = 383 Age > 60 Medical/Surgical	DOSS (treating physician) followed by electronic medical record		
Rudolph et al., 2016	Rudolph et al., 2016	External (temporal, geographic)	n = 246 Age > 55 Mixed	DSM-IV (researcher)	64 (26.0) any 43 (16.2) incident	0.69 (0.61-0.77) any 0.68 (0.59-0.77) incident
Adult medical inpatients						
Martinez et al., 2012	Martinez et al., 2012	External (temporal)	N = 302 Age > 18 Medical	CAM (researcher)	25.2 any	0.85 (0.85-0.88)
	Pendlebury et al., 2016	External (temporal, geographic)	N = 308 Age > 65 Medical	DSM IV (treating physician and researcher) following routine CAM screening	30.8 any 21.8 prevalent 9.1 incident	0.69 (0.62-0.76) any 0.62 (0.53-0.70) prevalent 0.78 (0.68-0.88) incident
Kobayashi et al., 2013		Internal (split sample)				
Older emergency department patients						
Han et al., 2009		none				
Hare et al., 2014		none				
Kennedy et al., 2014		Internal (boot-strapping)				0.77 (0.77-0.78)
Sri-On et al., 2016		none				
Adult inpatients in specific ward settings						
Patten et al., 2001		none				
Bannink et al., 2004		none				

(*continued on next page*)

Table 2 (*continued*)

Derivation study	Validation study	Validation type	Sample	Delirium measure	Delirium rate in validation, %	Discrimination AUC (95% confidence interval)
Neefjes et al., 2017		Internal (fivefold cross-validation)				0.65

common in studies of older medical and geriatric inpatients. Nine studies reported validation in a separate temporal and/or geographic cohort, and five tools were validated in separate studies.. Pendlebury (Pendlebury et al., 2016) undertook a within-subjects external validation of 3 tools included in this review (Douglas et al., 2013; Isfandiaty et al., 2012; J.A. Martinez et al., 2012) in a prospectively recruited cohort of 308 consecutive older patients. Brown et. al. (Brown et al., 2017) provided further validation of Douglas et al (Douglas et al., 2013). Rudolph (Rudolph et al., 2011) modified Inouye's tool (Inouye et al., 1993) to make it suitable to collect from medical record documentation, and Mestrez-Gonzalvo (Mestres Gonzalvo et al., 2017) validated deWit's predictive algorithm (de Wit et al., 2016). Most validation studies showed somewhat poorer performance in validation than derivation, with AUROC ranging from 0.64 to 0.85, and poorer discrimination for prevalent than incident delirium (Table 2).

3.6. Risk prediction performance

Thirteen studies suggested cut-points to identify low, intermediate and/or high risk groups (Bannink et al., 2004; Carrasco et al., 2014; Douglas et al., 2013; Han et al., 2009; Hare et al., 2014; Inouye & Charpentier, 1996; Inouye et al., 1993; Kennedy et al., 2014; J.A. Martinez et al., 2012; Pendlebury et al., 2017; Pompei et al., 1994; Rudolph et al., 2016; Tadiri, 2016), as outlined in Table S2 of the supplementary material. Only studies in older medical or medical/surgical patients considered more than one risk cut-point. Classification measures including sensitivity, specificity, and positive and negative predictive values are reported in Table 4. Studies could not be assessed if they were case-control design (de Wit et al., 2016; Levkoff et al., 1988; Patten et al., 2001), added extra delirium cases to increase disease prevalence post-hoc (Neefjes et al., 2017), did not suggest a specific cut-point to identify separate risk groups (Francis et al., 1990; O'Keeffe & Lavan, 1996), or provided insufficient information to calculate required values (Isfandiaty et al., 2012). Cut-offs between low and intermediate risk identified 22-90% patients as at risk, generally higher (60-90%) in older medical patient samples than in all age adult inpatients or emergency department samples (Table 4). Sensitivity was generally greater than 80%, with high negative predictive values in all settings. Specificity was poor in most studies, and positive predictive value was 15-46%. In studies which specified an additional high risk cut-point, 10-41% of participants were identified, and specificity was generally greater than 80% but sensitivity was poor.

4. Discussion

4.1. Summary of findings

This systematic review identified 22 prediction tools for delirium in adult non-surgical hospital inpatients published since 1988. Of these, fourteen have been externally validated, with most validation studies in older medical and geriatric patients. Most studies used a validated clinical assessment of delirium, although the largest studies were retrospective using electronic records and reported very low rates of delirium, consistent with poor recognition and documentation of delirium (Inouye, 2006). Tools used varying combinations of risk factors, with inconsistency in definitions and measurement of risk factors between studies. Nonetheless, there are some consistent findings, particularly amongst predisposing factors. Older age was generally a significant risk factor in studies which included participants aged less than 65, but less likely to be important in studies confined to older patients. Patients with cognitive and/or functional impairment would be classified at least as intermediate risk by most models, across the range of settings studied. Illness severity, infection, and biochemical abnormalities were studied across all settings, but inconsistent selection and definitions of variables likely contribute to inconsistent findings for these precipitating factors. The most common modelling approach was logistic regression, assuming an additive (linear) effect of included variables. No approaches to modelling appear to have built on Inouye's early finding of a "double gradient" effect between predisposing and precipitating factors (Inouye & Charpentier, 1996). Model discrimination was reported in most studies but calibration was poorly reported. A recently published review by Lindroth et al (Lindroth et al., 2018) provides detailed commentary on modelling, and recommends consideration of more advanced statistical techniques and better study reporting (Collins, Reitsma, Altman, & Moons, 2015).

4.2. Application to practice

Despite the variation in risk factors, populations sampled, and modelling approaches, most models reported fair to good discrimination. However, studies did not provide specific advice to clinicians about how to apply the tool in clinical practice. Our analysis of tool performance may provide some guidance. Most tools with a cut-off score discriminating between low and moderate to high risk were able to identify 80% or more of those who developed delirium. However in high-prevalence populations such as older medical patients the majority of patients would be considered at increased risk; in such circumstances it could be argued that resources might better be spent on delivering preventive interventions and monitoring for all patients, rather than diverting resources to screening. However in lower prevalence population such as younger and middle aged medical patients or emergency department populations, where only one third to one half of patients were flagged at risk, a screening tool might be useful for identifying a larger low risk group who do not require preventive interventions or monitoring for development of delirium. Although using a high risk cut-off reduced the proportion of identified patients to 10-40% of the sample, this strategy generally had poor sensitivity, so that many true cases would not be identified for preventive interventions at admission. This strategy might be more appropriate for selecting a small, high risk group for trials of new preventive interventions (e.g. new pharmacological agents) where there may be high costs or risk of side effects.

Studies of non-pharmacological interventions in non-surgical patients have not used consistent risk stratification approaches. Although Inouye's landmark Hospital Elder Life Program study (Inouye et al., 1999) used their 1993 tool (Inouye et al., 1993) for identification and targeting interventions, subsequent studies have used age alone (Rubin, Neal, Fenlon, Hassan, & Inouye, 2011; Young et al., 2015) or other validated or non-validated tools (Brown et al., 2018; Martinez, Tobar, Beddings, Vallejo, & Fuentes, 2012; Strijbos, Steunenberg, van der Mast, Inouye, & Schuurmans, 2013). No study we found has reported intervention effectiveness by risk level, and we found little evidence of translation into routine clinical practice (Siddiqi, 2016). A study reporting risk screening and flagging in the emergency department did not show any impact on subsequent delirium diagnosis or outcomes (Arendts et al., 2017).

Table 3
Delirium risk factors tested and included in prediction models. □ indicates that the factor was tested but not included in final model; ■ indicates that the factor was included in final model. Summary counts of the number of studies which tested each factor and which included each factor in their final model are provided in the first two rows.

	Predisposing factors									Precipitating factors							
	Age	Cognitive Impairment	Social Determinants	Drug	Function	Comorbidity	Sensory Impairment	Alcohol Use	Mood	Infection	Illness severity	↑ Urea and/ or Creatinine	Electrolytes	Low albumin	Iatrogenic interventions	New Medications	Pain
Tested (n)	21	19	17	16	12	13	12	10	8	19	18	14	13	7	7	2	1
Included (n)	10	16	0	4	6	4	6	2	2	7	9	4	5	3	3	2	0
Included if tested (%)	48	84	0	25	50	31	50	20	25	37	50	29	38	43	43	100	0
Older medical and geriatric inpatients																	
Francis et al., 1990	□	■	□	■	□					■	■	■	■		□		
Inouye et al., 1993	□	■	□		□		■		□		■	■					
Inouye and Charpentier, 1996										□	□	□		■	■	■	
O'Keeffe et al., 1996	□	■	□	□	□	□	□	□	□	□	■	■	□	□			
Isfandiaty et al., 2012	□	■	□	□	■	□				■	□	□	□	□			
Douglas et al., 2013	■	■	□	□	□	□	□	□	□	□	■	□	□				□
Carrasco et al., 2014	□	□	□		■	□				□	□	■	□	□			
Tadiri, 2016	■	■	□	□		■			□	□	■	□	■				
Pendlebury et al., 2017	■	■					■			■	■						
Older mixed hospital inpatients																	
Levkoff et al., 1988	□		□	□		□				■	□	□	□	■	□		
Pompei et al., 1994	□	■	□		□	■		■	■								
deWit 2016	■		□	■		□				□	□	□	□	□			
Rudolph et al., 2016	■	■					■			■	■						
Adult medical inpatients																	
Martinez et al., 2012	■	□	□	□	■		□	□		□	□	□	□		□	■	
Kobayashi et al., 2013	■	■	□	□	■	■	□	□	□	□							
Older emergency department patients																	
Han et al., 2009	□	■	□	□	■		■			□	□						
Hare et al., 2014	□	■	□	□	□	□	□	□	■	□	■						
Kennedy et al., 2014	■	■	□	□		□			□	■	■	□	□				
Sri-On et al., 2016	□	■	□	□			■	□		□	□	□	■				
Adult inpatients in specific ward settings																	
Patten et al., 2001	■	■	□	■		■		□							□		
Bannink et al., 2004	■	■		■			■	■		■			■		■		
Neefjes et al., 2017	□	□		□	■	□	□	□		□	□	□	■	■	■		

Table 4
Classification measures for different tools at cut-offs defined by authors of derivation and/or validation studies. Sens = sensitivity (the proportion of people with delirium who were identified at risk); spec = specificity (the proportion of people with delirium who were identified low risk); PPV = positive predictive value (the proportion of people identified at risk who developed delirium); NPV = negative predictive value (the proportion of people identified low risk who did not develop delirium).

			Intermediate risk cut-off					High risk cut-off				
Derivation study Validation(s)	Delirium type	Delirium rate, %	Identified at risk, %	sens	spec	PPV	NPV	Identified at risk, %	sens	spec	PPV	NPV
Older medical and geriatric inpatients												
Inouye et al., 1993	incident	25.4	68.9	.89	.38	.33	.91	11.3	.37	.98	.83	.82
Inouye et al., 1993		16.7	82.5	.97	.20	.20	.97	22.2	.41	.82	.32	.87
Pendlebury et al., 2016		9.1	70.3	.95	.34	.19	.98	5.5	.14	.96	.38	.87
Inouye et al., 1996	incident	17.9	61.2	.94	.46	.28	.97	11.2	.37	.94	.59	.87
Inouye et al., 1996		15.1	59.9	.89	.45	.23	.96	9.9	.23	.93	.36	.87
Rudolph et al., 2011		14.4	90.0	.92	.10	.15	.89	13.3	.23	.88	.25	.87
Douglas et al., 2013	incident	12.0	39.7	.79	.66	.25	.96					
Douglas et al., 2013		8.5	49.1	.79	.54	.14	.96					
Pendlebury et al., 2016		9.1	56.3	.95	.50	.24	.98					
Brown et al., 2017		5.8	21.9	.50	.89	.25	.96					
Carrasco et al., 2014	incident	6.7	30.2	.88	.74	.20	.99					
Tadiri, 2016	incident	14.8	58.0	.82	.54	.24	.95					
Pendlebury et al., 2017	any	30.8	85.4	.95	.19	.41	.86	28.3	.57	.88	.74	.78
Older mixed hospital inpatients												
Pompei et al., 1994	any	14.8	60.9	.92	.45	.22	.97	21.3	.47	.83	.33	.90
Pompei et al., 1994		26.5	58.3	.83	.50	.38	.89	20.1	.35	.85	.46	.78
Rudolph et al., 2016	Prevalent	8.5	83.1	.97	.18	.10	.99	38.3	.79	.66	.18	.97
Rudolph et al., 2016	any	26.0	58.9	.77	.47	.34	.85	41.1	.64	.67	.41	.84
Adult medical inpatients												
Martinez et al., 2012	any	13.1	44.8	.83	.61	.24	.96					
Martinez et al., 2012		25.2	52.9	.93	.61	.44	.97					
Pendlebury et al., 2016		30.8	73.8	.90	.36	.46	.85					
Older emergency department patients												
Han et al., 2009	prevalent	8.3	54.5	.96	.49	.15	.99					
Hare et al., 2014	prevalent	7.2	33.8	.87	.70	.19	.96					
Kennedy et al., 2014	prevalent	9.3	38.5	.76	.65	.19	.96					
Adult inpatients in specific ward settings												
Bannink et al., 2004	any	12.5	49.0	.92	.57	.24	.98					

4.3. Strengths and limitations

Strengths of our review include prospective registration, a broadly inclusive search strategy, detailed assessment of bias using a tool tailored for predictive studies (Moons et al., 2014) and detailed analysis of risk factor patterns and tool classification which provides greater clarity about the potential implications for tool use in clinical practice. Our findings extend the results of previous reviews. Newman (Newman et al., 2015) reviewed delirium risk stratification models in adult inpatients, but was limited to studies with external validation. Like us, the authors highlighted the range and inconsistency in risk factor collection between studies, and a lack of studies addressing clinical application of the tools. Lindroth (Lindroth et al., 2018)included medical and surgical inpatient groups, and provided detailed review of modelling methods, but included only older patients and excluded studies in the emergency department, which may limit broad clinical application. Limitations of our review include exclusion of non-English studies, and lack of data for reporting calibration and classification measures for all studies. In addition, the lack of consistency between tools applied in different clinical settings, and the few validation studies outside older medical patients, means that we cannot recommend a single “best” tool to apply across different settings.

4.4. Conclusions

In summary, several risk screening tools are available which may predict risk of delirium in medical wards, emergency departments and specialist medical settings with reasonable discrimination. Whilst domains such as cognitive and functional impairment appear important across settings, high variability in how these domains were measured highlights the lack of consistent approaches to screening these factors, hampering practical application of tools and precluding clear recommendations about the best approach to risk screening. The value of applying screening tools will depend on risk factor prevalence in the target population, and the availability and cost of preventive interventions. Importantly, screening needs to be clearly linked to evidence-based actions. Delirium prevention and intervention studies applying one or more of these validated risk tools to report benefits for different levels of risk would provide practical information about the value of risk stratification in targeting preventive interventions.

Conflict ofintereststatement

Nothing decalred.

Acknowledgements

This work was supported by a Queensland Health and Medical Research Junior Doctor Fellowship awarded to Dylan Flaws. The funders had no role in undertaking research, manuscript preparation or decision to submit. We thank University of Queensland Medical Librarian Lars Eriksson for assistance with search terms.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.archger.2019.05.013.

References

Arendts, G., Love, J., Ngree, Y., Burce, D., Hare, M., & Dey, I. (2017). Rates of delirium diagnosis do not improve with delirium risk screening: results of the Emergency

Department Delirium Initiative trial. *Journal of the American Geriatrics Society, 65*(8), 1810–1815.
Bannink, M., van Veluw, H., van Zuijlen, L., Enting, R., & van der Rijt, K. (2004). Risk factors predicting the development of delirium. *European Journal of Palliative Care, 11*(6), 239–240.
Brown, E. G., Josephson, S. A., Anderson, N., Reid, M., Lee, M., & Douglas, V. C. (2017). Predicting inpatient delirium: The AWOL delirium risk-stratification score in clinical practice. *Geriatric Nursing*.
Brown, E. G., Josephson, S. A., Anderson, N., Reid, M., Lee, M., & Douglas, V. C. (2018). Evaluation of a multicomponent pathway to address inpatient delirium on a neurosciences ward. *BMC Health Services Research, 18*, 106.
Carrasco, M. P., Villarroel, L., Andrade, M., Calderón, J., & González, M. (2014). Development and validation of a delirium predictive score in older people. *Age & Ageing, 43*(3), 346–351.
Cole, M. G., McCusker, J., Bellavance, F., Primeau, F. J., Bailey, R. F., Bonnycastle, M. J., ... Laplante, J. (2002). Systematic detection and multidisciplinary care of delirium in older medical inpatients: a randomized trial. *Canadian Medical Association Journal, 167*(7), 753–759.
Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ, 350*, g7594.
de Wit, H. A. J. M., Winkens, B., Mestres Gonzalvo, C., Hurkens, K. P. G. M., Mulder, W. J., Janknegt, R., ... Schols, J. M. G. A. (2016). The development of an automated ward independent deliriumrisk prediction model. *International Journal of Clinical Pharmacy, 38*(4), 915–923.
Douglas, V. C., Hessler, C. S., Dhaliwal, G., Betjemann, J. P., Fukuda, K. A., Alameddine, L. R., ... Josephson, S. A. (2013). The AWOL tool: Derivation and validation of a delirium prediction rule. *Journal of Hospital Medicine, 8*(9), 493–499.
Eeles, E. M. P., Hubbard, R. E., White, S. V., O'Mahony, M. S., Savva, G. M., & Bayer, A. J. (2010). Hospital use, institutionalisation and mortality associated with delirium. *Age & Ageing, 39*(4), 470–475.
Elie, M., Cole, M. G., Primeau, F. J., & Bellavance, F. (1998). Delirium risk factors in elderly hospitalized patients. *Journal of General Internal Medicine, 13*, 204–212.
Francis, J., Martin, D., & Kapoor, W. N. (1990). A prospective study of delirium in hospitalized elderly. *Journal of the American Medical Association, 263*(8), 1097–1101.
Han, J. H., Zimmerman, E. E., Cutler, N., Schnelle, J., Morandi, A., Dittus, R. S., ... Ely, E. W. (2009). Delirium in older emergency department patients: recognition, risk factors, and psychomotor subtypes. *Academic Emergency Medicine, 16*(3), 193–200.
Hare, M., Arendts, G., Wynaden, D., & Leslie, G. (2014). Nurse screening for delirium in older patients attending the emergency department. *Psychosomatics, 55*(3), 235–242.
Holden, J., Jayathissa, S., & Young, G. (2008). Delirium among elderly general medical patients in a New Zealand hospital. *Internal Medicine Journal, 38*(8), 629–634.
Inouye, S. K. (1998). Delirium in hospitalized older patients: Recognition and risk factors. *Journal of Geriatric Psychiatry and Neurology, 11*(3), 118–125.
Inouye, S. K. (2006). Delirium in older persons. *New England Journal of Medicine, 354*(11), 1157–1165.
Inouye, S. K., Bogardus, S. T., Charpentier, P., Leo-Summers, L., Acampora, D., Holford, T. R., ... Cooney, L. M. (1999). A multicomponent intervention to prevent delirium. *New England Journal of Medicine, 340*, 669–676.
Inouye, S. K., & Charpentier, P. A. (1996). Precipitating factors for delirium in hospitalized elderly persons: Predictive model and interrelationship with baseline vulnerability. *Journal of the American Medical Association, 275*(11), 852–857.
Inouye, S. K., Rushing, J. T., Foreman, M. D., Palmer, R. M., & Pompei, P. (1998). Does delirium contribute to poor hospital outcomes? *Journal of General Internal Medicine, 13*(4), 234–242.
Inouye, S. K., Viscoli, C. M., Horwitz, R. I., Hurst, L. D., & Tinetti, M. E. (1993). A predictive model for delirium in hospitalized elderly medical patients based on admission characteristics. *Annals of Internal Medicine, 119*(6), 474–481.
Iseli, R. K., Brand, C. A., Telford, M., & LoGiudice, D. (2007). Delirium in elderly general medical inpatients: a prospective study. *Internal Medicine Journal, 37*(12), 806–811.
Isfandiaty, R., Harimurti, K., Setiati, S., & Roosheroe, A. G. (2012). Incidence and predictors for delirium in hospitalized elderly patients: a retrospective cohort study. *Acta medica Indonesiana, 44*(4), 290–297.
Kennedy, M., Enander, R. A., Tadiri, S. P., Wolfe, R. E., Shapiro, N. I., & Marcantonio, E. R. (2014). Delirium risk prediction, healthcare use and mortality of elderly adults in the emergency department. *Journal of the American Geriatrics Society, 62*(3), 462–469.
Kobayashi, D., Takahashi, O., Arioka, H., Koga, S., & Fukui, T. (2013). A prediction rule for the development of delirium among patients in medical wards: Chi-square automatic interaction detector (CHAID) decision tree analysis model. *American Journal of Geriatric Psychiatry, 21*(10), 957–962.
Levkoff, S. E., Safran, C., Cleary, P. D., Gallop, J., & Phillips, R. S. (1988). Identification of factors associated with the diagnosis of delirium in elderly hospitalized patients. *Journal of the American Geriatrics Society, 36*(12), 1099–1104.
Lindroth, H., Bratzke, L., Purvis, S., Brown, R., Coburn, M., Mrkobrada, M., ... Sanders, R. (2018). Systematic review of prediction models for delirium in the older adult inpatient. *BMJ Open, 2018*(8), e019223.
Lundström, M., Edlund, A., Karlsson, S., Brännström, B., Bucht, G., & Gustafson, Y. (2005). A multifactorial intervention program reduces the duration of delirium, length of hospitalization, and mortality in delirious patients. *Journal of the American Geriatrics Society, 53*(4), 622–628.
Marcantonio, E. R. (2017). Delirium in hospitalized older adults. *New England Journal of Medicine, 377*(15), 1456–1466.
Martinez, F. T., Tobar, C., Beddings, C. I., Vallejo, G., & Fuentes, P. (2012). Preventing delirium in an acute hospital using a non-pharmacological intervention. *Age & Ageing, 41*, 629–634.
Martinez, J. A., Belastegui, A., Basabe, I., Goicoechea, X., Aguirre, C., Lizeaga, N., ... Emparanza, J. I. (2012). Derivation and validation of a clinical prediction rule for delirium in patients admitted to a medical ward: An observational study. *BMJ Open, 2*(5).
Mestres Gonzalvo, C., de Wit, H., van Oijen, B. P. C., Deben, D. S., Hurkens, K., Mulder, W. J., & van der Kuy, P. M. (2017). Validation of an automated delirium prediction model (DElirium MOdel (DEMO)): an observational study. *BMJ Open, 7*(11), e016654.
Moons, K. G. M., de Groot, J. A. H., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., ... Collins, G. S. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med, 11*(10), e1001744.
Neefjes, E. C. W., van der Vorst, M., Verdegaal, B., Beekman, A. T. F., Berkhof, J., & Verheul, H. M. W. (2017). Identification of patients with cancer with a high risk to develop delirium. *Cancer Medicine, 6*(8), 1861–1870.
Newman, M. W., O''Dwyer, L. C., & Rosenthal, L. (2015). Predicting delirium: a review of risk-stratification models. *General Hospital Psychiatry, 37*, 408–413.
O'Keeffe, S. T., & Lavan, J. N. (1996). Predicting delirium in elderly patients: development and validation of a risk-stratification model. *Age & Ageing, 25*(4), 317–321.
O'Keeffe, S. T., & Lavan, J. N. (1997). The prognostic significance of delirium in older hospital patients. *Journal of the American Geriatrics Society, 45*(2), 174–178.
Patten, S. B., Williams, J. V. A., Petcu, R., & Oldfield, R. (2001). Delirium in psychiatric inpatients: A case-control study. *Canadian Journal of Psychiatry, 46*(2), 162–166.
Pendlebury, S. T., Lovett, N., Smith, S. C., Cornish, E., Mehta, Z., & Rothwell, P. M. (2016). Delirium risk stratification in consecutive unselected admissions to acute medicine: Validation of externally derived risk scores. *Age & Ageing, 45*(1), 60–65.
Pendlebury, S. T., Lovett, N. G., Smith, S. C., Wharton, R., & Rothwell, P. M. (2017). Delirium risk stratification in consecutive unselected admissions to acute medicine: validation of a susceptibility score based on factors identified externally in pooled data for use at entry to the acute care pathway. *Age & Ageing, 46*, 226–231.
Pompei, P., Foreman, M., Rudberg, M. A., Inouye, S. K., Braund, V., & Cassel, C. K. (1994). Delirium in hospitalized older persons: Outcomes and predictors. *Journal of the American Geriatrics Society, 42*(8), 809–815.
Rubin, F. H., Neal, K., Fenlon, K., Hassan, S., & Inouye, S. K. (2011). Sustainability and scalability of the Hospital Elder Life Program at a community hospital. *Journal of the American Geriatrics Society, 59*, 359–365.
Rudolph, J. L., Doherty, K., Kelly, B., Driver, J. A., & Archambault, E. (2016). Validation of a Delirium Risk Assessment Using Electronic Medical Record Information. *Journal of the American Medical Directors Association, 17*(3), 244–248.
Rudolph, J. L., Harrington, M. B., Lucatorto, M. A., Chester, J. G., Francis, J., & Shay, K. J. (2011). Validation of a Medical Record-Based Delirium Risk Assessment. *Journal of the American Geriatrics Society, 59*, S289–294.
Siddiqi, N. (2016). Predicting delirium: Time to use delirium risk scores in routine practice? *Age & Ageing, 45*(1), 9–10.
Sri-On, J., Tirrell, G. P., Vanichkulbodee, A., Niruntarai, S., & Liu, S. W. (2016). The prevalence, risk factors and short-term outcomes of delirium in Thai elderly emergency department patients. *Emergency Medicine Journal, 33*(1), 17–22.
Strijbos, M. J., Steunenberg, B., van der Mast, R. C., Inouye, S. K., & Schuurmans, M. J. (2013). Design and methods of the Hospital Elder Life Program (HELP), a multicomponent targeted intervention to prevent delirium in hospitalized older patients: efficacy and cost-effectiveness in Dutch health care. *BMC Geriatrics, 13*, 78.
Tadiri, S. P. (2016). *Analyzing risk factors, resource utilization, and health outcomes of hospital-acquired delirium In elderly emergency department patients.* Boston University.
Tropea, J., Slee, J. A., Brand, C. A., Gray, L., & Snell, T. (2008). Clinical practice guidelines for the management of delirium in older people in Australia. *Australasian Journal on Ageing, 27*(3), 150–156.
Witlox, J., Eurelings, L. S. M., de Jonghe, J. F. M., Kalisvaart, K. J., Eikelenboom, P., & Van Gool, W. A. (2010). Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *Journal of the Amedican Medical Association, 304*(4), 443–451.
Young, J., Cheater, F., Collinson, M., Fletcher, M., Forster, A., Godfrey, M., ... Farrin, A. J. (2015). Prevention of delirium (POD) for older people in hospital: study protocol for a randomised controlled feasibility trial. *Trials, 16*, 340.
Young, J., Murthy, L., Westby, M., Akunne, A., & O'Mahony, R. (2010). Diagnosis, prevention, and management of delirium: summary of NICE guidance. *BMJ: British Medical Journal, 341*.