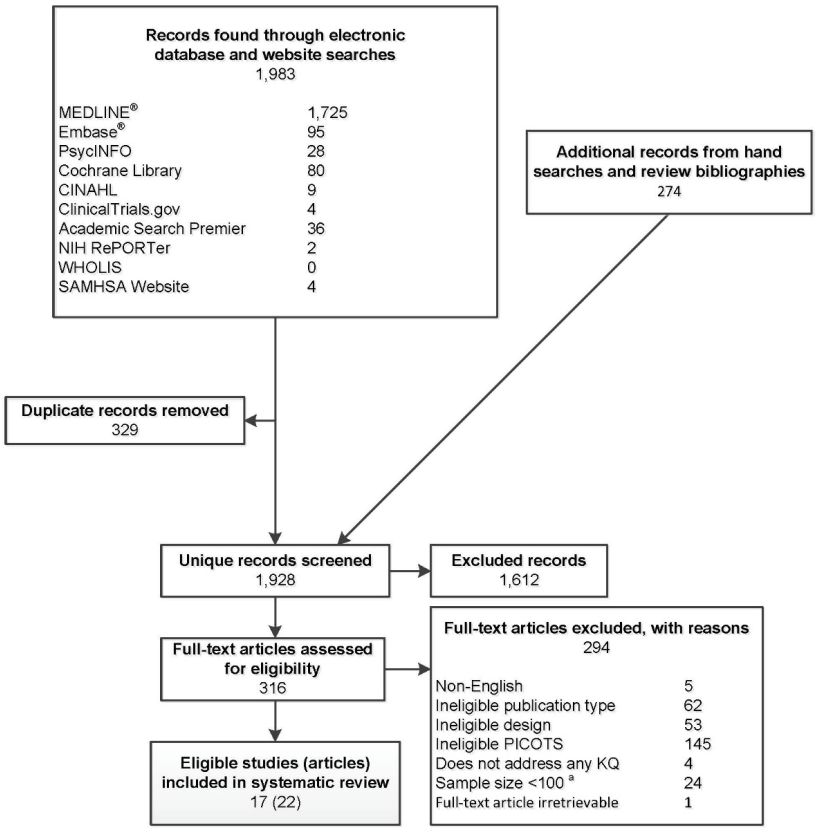
Figure 1. Summary of evidence search and selection of articles about strategies to prevent or de-escalate aggressive behavior



[a] This minimum sample size requirement only applies to nonrandomized studies.

CINAHL = Cumulative Index to Nursing and Allied Health Literature; KQ = Key Question; NIH RePORTer = National Institutes of Health Research Portfolio Online Reporting Tools; PICOTS = Populations-Interventions-Comparators-Outcomes-Time Frames-Settings; SAMHSA = Substance Abuse and Mental Health Services Administration; WHOLIS = World Health Organization's Library Database

Online Supplement Table 1. Eligibility criteria for review of strategies to de-escalate aggressive behavior

PICOTS	Inclusion	Exclusion
Populations	KQs 1 through 3: • Adult individuals (ages 18 or older) with an identified psychiatric disorder (if in an inpatient setting), including substance use disorders and delirium (but not dementia), or with severe psychiatric symptomatology (if in an emergency department setting where a formal psychiatric diagnosis often is not made), who are at risk of or actively exhibiting aggressive behavior toward self, others, or property.	All other populations
Interventions	KQs 1a and 2a: • Strategies (early intervention techniques) targeted to reduce the likelihood of aggressive behavior (examples provided in the PICOTS criteria) KQs 1b/1c and 2b/2c: • Strategies targeted to decrease aggression for those who are actively aggressive (examples provided in the PICOTS criteria) KQ 3: Same as KQs 1 and 2	All other interventions • For medication-based interventions, those that are not FDA-approved for any indication
Comparators	KQs 1a and 2a: • Other strategies (early intervention techniques), but not seclusion and restraints, targeted to reduce the likelihood of aggressive behavior, as described above for KQs 1a and 2a • Usual care, defined as the standard of care for a particular setting before implementation of an intervention designed to decrease the likelihood of aggression and/or the use of seclusion and restraint KQs 1b/1c and 2b/2c: • Other strategies targeted to decrease aggression for those who are actively aggressive, as described above for KQs 1b/1c and 2b/2c • Seclusion or restraint (for 1b and 2b only) (as defined in the PICOTS criteria) • Usual care, defined as the standard of care for a particular setting before implementation of an intervention designed to decrease aggression and/or the use of seclusion and restraint KQ 3: Same as KQs 1 and 2	All KQs: • A study with no comparison group • For medication-based strategies, placebo-only comparisons and those comparing different doses or routes of administration
Outcomes	KQs 1a, 1b, and 1c: • Intermediate outcomes: – Primary outcomes: o Decreased aggression in terms of frequency, severity, or duration (as measured by direct counts or by validated aggression scales) o KQs 1a and 1c only: Reduced use of seclusion or restraints (decreased rate, amount, or duration) o To be eligible, each study must have reported on at least one of the outcomes above – Secondary outcomes: o As defined in the PICOTS criteria • Final health outcomes: – As defined in the PICOTS criteria KQs 2a, 2b, and 2c: As defined in the PICOTS criteria KQ 3: Same as KQs 1 and 2	None

Online Supplement Table 1. Eligibility criteria for review of strategies to de-escalate aggressive behavior (continued)

PICOTS	Inclusion	Exclusion
Timing	All KQs: Imminently or within current episode of care (e.g., inpatient hospitalization, emergency department stay)	All KQs: Outside current episode of care
Settings	All KQs: Acute care settings, including emergency department or hospital (e.g., private or public psychiatric hospitals, general medical hospitals at which discharge occurs within 35 days of beginning treatment)[a]	All KQs: Outpatient, community-based, jails, prisons, schools, chronic care, forensic-only,[b] or long-term care settings
Study designs	All KQs: • Systematic reviews, with or without meta-analyses • Randomized controlled trials • Nonrandomized controlled trials • Cohorts (prospective and retrospective) • Case-control studies • Single group pre/post studies (including pre/post studies with <3 pre- and <3 post-intervention time points)[c, d] • Interrupted time-series designs (i.e., time-series studies with ≥3 pre-intervention and ≥3 post-intervention measurements with one or more groups)[c]	All KQs: • Case studies or series • Cross-sectional studies • Studies without a comparison group • Nonsystematic review
Publications	All KQs: Original research	All KQs: Not original research (e.g., editorials without original data, newspaper articles)
Geographic locations	Developed countries ("very high" human development index per the United Nations Development Programme[1])	All other countries
Language	English	All other languages

[a] Studies of settings that treated patients receiving both acute and chronic care were excluded. To be clear, a single unit or wing of a hospital could be eligible if inpatient stays were 35 days or less, even if other sections of the larger hospital provided longer-term care. We assumed that studies describing their sample's inpatient clinical services as "acute care" referred to discharge within 35 days of admission, when no specific information about lengths of stay was available. We attempted to locate information about the types of care provided in study-specific settings if there was concern that study analyses may have included a mixture of acute-care and chronic-care patients. When no information was available to confirm that a study's inpatient clinical services were acute care or that lengths of inpatient stays were 35 days or less, we excluded it.

[b] We excluded studies focusing only on forensic units or hospitals, but studies conducted in acute care settings were eligible if their samples included both forensic and nonforensic patients.

[c] A "group" could indicate a group of patients, acute care unit, or hospital evaluated before and after implementation of an intervention.

[d] We considered time-series studies with 2 pre-intervention and/or 2 post-intervention measurements as pre/post studies.

FDA = U.S. Food and Drug Administration; KQ = Key Question; PICOTS = populations, interventions, comparators, outcomes, timing, and settings.

References

1. United Nations Development Programme (UNDP). Human Development Report 2014 - Sustaining Human Progress: Reducing Vulnerabilities and Building Resilience. United Nations Development Programme; 2014 http://hdr.undp.org/en/2014-report.

Online Supplement Table 2. Key characteristics of studies of interventions to de-escalate aggressive behaviors in acute care settings (including psychiatric diagnoses and sociodemographic characteristics), by intervention category

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
Staff Training Interventions					
Kontio et al., 2014[1] CRT, High Psychiatric hospitals (8 units), Finland	NR 2 years	G1: Online eLearning course for unit nurses on managing aggression or violence and preventing coercion G2: Education as usual	Inpatients on acute, closed units that practice seclusion or restraint	NR	NR
Smoot et al., 1995[2] CRT, High Inpatient psychiatric recidivist units[c], United States	NR[b] 6 months	G1: Empathic interpersonal communication training program for hospital staff G2: Usual care	Primary diagnosis of mental illness for patients who had returned to the hospital within 1 year of a previous discharge	NR	NR
Risk Assessment Interventions					
Abderhalden et al., 2008[3] CRT, Medium Psychiatric inpatient treatment facilities, Switzerland	973[d] 3 months	G1: Structured risk assessment (BVC) for every new patient twice a day during the first 3 days of hospitalization (n=390) G2: Usual care (n=583)	Inpatients, most with an acute psychiatric disorder	Schizophrenia, schizotypal and delusional disorders G1: 33.4 G2: 35.7 Disorders due to psychoactive substance use G1: 26.2 G2: 24.2 Mood (affective) disorders G1: 15.5 G2: 15.3 Neurotic, stress-related and somatoform disorders, behavioral syndromes associated	Age G1: 39.0 (13.1) G2: 38.0 (14.3) Percent female G1: 45.6 G2: 44.8 Percent non-white: NR

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
				with physiological disturbances and physical factors G1: 14.3 G2: 11.5 Other/missing G1: 10.6 G2: 13.3	
van de Sande et al., 2011[4] CRT, Medium Acute psychiatric units, Netherlands	458 30 weeks	G1: Structured risk assessment (n=207): Daily (5 mins) using BVC and Kennedy-Axis V (short version); Weekly (15 mins) using Kennedy-Axis V (full version), BPRS, Dangerousness Scale, and the SDAS G2: Usual care / Treatment as usual (n=251)	Patients admitted to acute psychiatric units, mostly with psychotic disorders (74%) and personality disorders (25%)	Patients admitted to acute psychiatric units Psychotic disorder G1: 74 G2: 57 Personality disorders G1: 25 G2: 6 Drug misuse first diagnosis G1: 4 G2: 3	Age (SD) G1: 38 (13) G2: 40 (11) Percent female G1: 47 G2: 46 Percent non-white G1: 31 G2: 16
Multimodal Interventions					
Putkonen et al., 2013[5] CRT, Medium Public psychiatric hospital, Finland	NR[e] 6 months	G1: Six Core Strategies implementation (best practices to reduce use of seclusion and restraints) G2: Usual care / Treatment as usual	Male inpatients in high-security units who had psychotic illness and a history of violence	Schizophrenia: 100	Age G1: 40.2 (10.6) G2: 38.4 (10.6) Percent female: 0 Percent non-white: NR
Environmental or Group Psychotherapeutic Interventions					

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
Nurenberg et al., 2015[6] RCT, Medium State psychiatric hospital, United States	90 3 months	G1: Equine-assisted psychotherapy (n=24) G2: Canine-assisted psychotherapy (n=25) G3: Environmentally enhanced social skills group psychotherapy (n=23) G4: Usual care (n=18)	Inpatients with “aggressive or regressed behavior” or ”persistent social isolation” and difficulty engaging in discharge-related programs	Schizophrenia G1: 29 G2: 32 G3: 35 G4: 39 Schizoaffective G1: 38 G2: 56 G3: 43 G4: 28 Affective/Other G1: 33 G2: 12 G3: 22 G4: 33	Age G1: 44.3 (13.8) G2: 45.0 (10.8) G3: 43.2 (10.3) G4: 44.4 (11.9) Female G1: 25 G2: 44 G3: 43 G4: 33 Non-White G1: 38 G2: 32 G3: 43 G4: 45
Carlson et al., 1993[7] Retrospective cohort study, High State psychiatric hospital, United States	120 90 days	G1: Occupational therapy at least 1 time every 30 days (n=60) G2: No occupational therapy in at least 1 of the 3 30-day periods (n=60)	Patients with at least a 90-day inpatient stay on psychiatric unit; only data from first 90 days of stay were included	Schizoaffective disorder G1: 16.7 G2: 21.7 Schizophrenia, paranoid type, chronic G1: 16.7 G2: 18.4	Age G1: 47.6 (18.3) G2: 45.5 (15.8) Female G1: 52 G2: 48 Non-White G1: 19 G2: 15
Medication Protocols					
Bieniek et al., 1998[8] RCT, Low Psychiatric emergency service (in-hospital), United States	20 3 hours	G1: Haloperidol, 5 mg i.m. plus lorazepam 2 mg i.m. (n=9) G2: Lorazepam, 2 mg i.m. (n=11)	Patients with serious, acutely agitated or aggressive behavior and who met clinical criteria for use of chemical restraints	Bipolar disorder, manic G1: 33.3 G2: 54.5 Psychosis NOS	Age Overall (mean, SD): 36.3 (8.1) G1 (median): 41.0 G2 (median): 35.0

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
				G1: 22.2 G2: 18.2 Schizophrenia, paranoid G1: 11.1 G2: 18.2 Brief reactive psychosis G1: 11.1 G2: 0 Schizophrenia, undifferentiated G1: 11.1 G2: 0 Substance-induced G1: 11.1 G2: 9.1	Female G1: 44.4 G2: 27.3 Hispanic G1: 33.3 G2: 18.2 African-American G1: 33.3 G2: 54.5 Haitian G1: 0 G2: 9.1
Dorevitch et al., 1999[9] RCT, Medium Psychiatric hospital, Israel	28 90 minutes	During aggressive event: G1: Haloperidol, 5 mg i.m. (n=13) G2: Flunitrazepam, 1 mg i.m. (n=15)	Acute unit patients with active psychosis, disruptive or aggressive behavior, pronounced psychomotor agitation or violent outbursts	Schizophrenia & schizoaffective disorder G1: 92.3 G2: 93.3 Bipolar I disorder G1: 7.7 G2: 6.7	Age G1: 36.8 (15.1) G2: 34.9 (8.1) Female G1: 61.5 G2: 46.7 Non-white: NR
Georgieva et al., 2013[10] RCT, High Psychiatric hospital, Netherlands	520 (with 659 admissions) 144 weeks	Intervention of first choice for agitation and risk of violence: G1: Involuntary medication (n=236, with 306 admissions)	Patients admitted to acute units, most with either addiction or a psychotic, mood, personality, or post-traumatic stress disorder	Psychotic disorder G1: 20 G2: 20 Mood disorder G1: 31	Age G1: 40 (13) G2: 40 (12) Female G1: 52

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
		G2: Seclusion (n=284, with 353 admissions)		G2: 32 Personality disorder G1: 24 G2: 23 Addiction G1: 31 G2: 32 PTSD G1: 5 G2: 8	G2: 47 Non-Dutch ethnicity G1: 17 G2: 18
Isbister et al., 2010[11] RCT, Medium Public psychiatric hospital, Australia	91 6 hours	G1: Droperidol, 10 mg i.m. (n=33) G2: Midazolam, 10 mg i.m. (n=29) G3: Droperidol, 5 mg i.m. plus midazolam, 5 mg i.m. (n=29)	Patients presenting to the emergency department with violence and acute behavioral disturbance and requiring both physical restraint and parenteral sedation	Alcohol intoxication G1: 70 G2: 76 G3: 66 Deliberate self-harm G1: 48 G2: 41 G3: 45 Drug-induced delirium G1: 6 G2: 10 G3: 10 Acute psychosis G1: 6 G2: 3 G3: 6 Other	Age, mean (range) G1: 37 (25 to 45) G2: 35 (27 to 43) G3: 30 (22 to 40) Female G1: 64 G2: 38 G3: 48 Non-white: NR

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
				G1: 3 G2: 0 G3: 3	
Krakowski et al., 2006[12; 13] RCT, Medium State psychiatric facilities (in-hospital), United States	110 12 weeks	G1: Clozapine, oral 500 mg/day (n=37) G2: Olanzapine, oral 20 mg/day (n=37) G3: Haloperidol, oral 20 mg/day (n=36)	Patients with confirmed episode of physical assault directed at another person during their current hospitalization and some persistence of aggression	Schizophrenia G1: 73 G2: 62.2 G3: 58.3 Schizoaffective disorder G1: 27 G2: 37.8 G3: 41.7	Age G1: 35.1 (12.3) G2: 35.6 (9.4) G3: 32.7 (10.6) Female G1: 16.2 G2: 21.6 G3: 16.7 Black G1: 54.1 G2: 75.7 G3: 58.3 Hispanic G1: 21.6 G2: 10.8 G3: 22.2 Other G1: 5.4 G2: 0 G3: 0
Michaud et al., 2014[14] Retrospective cohort study, High Public psychiatric hospital, U.S.	200 24 hours	G1: Delirium treatment within 24 hours (n=102) G2: No delirium treatment, or treatment after 24 hours (n=98)	Adults in an intensive care unit with a documented positive delirium screen at time of mechanical ventilation	NR	Age G1: 58 (17) G2: 62 (15) Female G1: 53 G2: 53

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
					Non-white: NR
Richards et al., 1998[15] RCT, High Large urban university emergency department, United States	202 60 minutes	G1: Droperidol, 2.5-5 mg i.v.[f] (n=102) G2: Lorazepam, 2-4 mg i.v.[f] (n=100)	Acutely agitated patients with violent, controlled, or uncontrolled muscular movement placing themselves and staff at danger and requiring constant supervision	NR, but toxicology tests positive for following substances: Methamphetamine G1: 70.6 G2: 74.0 Cocaine G1: 15.7 G2: 12.0 Ethanol G1: 49.0 G2: 48.0	Age Overall: 33.9 (10.5) G1: 33.2 (10.2) G2: 34.6 (10.8) Percent female Overall: 38.1 G1: 39.2 G2: 37.0 Percent non-white Overall: 30.7 G1: 31.4 G2: 30
Villari et al., 2008[16] NRCT, Medium Psychiatric emergency service (in-hospital), Italy	101 72 hours	G1: Risperidone, oral 2-6 mg/day (n=27) G2: Olanzapine, oral 10-20 mg/day (n=24) G3: Quetiapine, oral 300-800 mg/day (n=22) G4: Haloperidol, oral 5-15 mg/day (n=28)	Psychotic inpatients requiring emergency medication for control of agitation	Schizophrenia G1: 30 G2: 46 G3: 32 G4: 40 Schizoaffective disorder G1: 7 G2: 0 G3: 27 G4: 11 Brief psychotic disorder G1: 48 G2: 16 G3: 18 G4: 32	Age G1: 39.2 (12.7) G2: 41.5 (12.2) G3: 41.2 (15.2) G4: 39.8 (9.0) Female G1: 44.4 G2: 37.5 G3: 59.1 G4: 39.3 Percent non-white: NR

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
				Delusional disorder G1: 15 G2: 27 G3: 14 G4: 7 Bipolar I disorder G1: 0 G2: 21 G3: 9 G4: 11	
Volavka et al., 2004[17; 18] RCT, Medium State psychiatric hospitals, United States	157 14 weeks	G1: Clozapine, oral 500 mg/day (n=40) G2: Olanzapine, oral 20 mg/day (n=39) G3: Risperidone, oral 8 mg/day (n=41) G4: Haloperidol, oral 20 mg/day (n=37)	Treatment-resistant inpatients diagnosed with chronic schizophrenia or schizoaffective disorder	Schizophrenia: 86 Schizoaffective disorder: 14	Age: 40.8 (9.2)
Wilhelm et al., 2008[19] NRCT, High Psychiatric or forensic hospitals (n=102), Germany	558 6 days[g]	G1: Olanzapine, oral dose NR[h] (n=390) G2: Non-olanzapine medication, oral dose NR[h] (n=168) G3: Risperidone, oral dose NR[h] (n=72) G4: Non-risperidone medication, oral dose NR[h] (n=486) G5: Haloperidol, oral dose NR[h] (n=132) G6: Non-haloperidol medication, oral dose	Inpatients newly admitted to a psychiatric (98%) or forensic hospital (2%) with psychiatric disorders who presented with agitation with or without aggression and required antipsychotic treatment	Primary psychiatric diagnoses[i] Schizophrenia spectrum disorders Overall: 59.1 G1: 55.1 vs. G2: 68.5 G3: 69.4 vs. G4: 57.6 G5: 69.7 vs. G6: 55.9 Substance use disorders G1: 17.7 vs. G2: 17.3 G3: 9.7 vs. G4: 18.7 G5: 17.4 vs. G6: 17.6	Age, median (range) Overall: 38 (18 to 93) G1: 37 (18 to 93) G2: 39 (19 to 84) G3: 40 (19 to 87) G4: 38 (18 to 93) G5: 39 (18 to 93) G6: 38 (18 to 90) Percent female Overall: 36.7 G1: 39.2

Author, Year Study Design, Risk of Bias Clinical Setting, Country	N of Patients[a] Duration of Intervention(s)	Intervention(s) and Comparator(s) (n of patients, if reported)	Summary Description of Patient Population	Percent with Psychiatric Diagnoses	Age: Mean (SD) Percent Female Percent Non-White
		NR[h] (n=426)		Mood (affective) disorders G1: 20.5 vs. G2: 4.8 G3: 5.6 vs. G4: 17.3 G5: 11.4 vs. G6: 17.1 Adult personality and behavior disorders G1: 17.2 vs. G2: 10.1 G3: 13.9 vs. G4: 15.2 G5: 3.0 vs. G6: 18.8 Organic disorders, including symptomatic mental disorders G1: 10.0 vs. G2: 17.9 G3: 19.4 vs. G4: 11.3 G5: 14.4 vs. G6: 11.7 Other disorders[j] G1: 11.0 vs. G2: 8.3 G3: 11.1 vs. G4: 10.1 G5: 3.8 vs. G6: 12.2	G2: 31.0 G3: 36.1 G4: 36.8 G5: 29.5 G6: 39.0 Percent non-white: NR

[a] The number of patients reflects the entire study from baseline through post-intervention or longer-term followup.
[b] Average of 92 patients discharged per month in each unit, meaning about 184 patients were included in the study each month.[2]
[c] The two study units specialized in caring for people with a primary diagnosis of mental illness who had returned to the hospital within 1 year of a prior discharge.[2]
[d] Neither the baseline nor intervention period count includes patients admitted to the five units that preferred to introduce the study protocol of structured risk assessment without randomization.[3]
[e] Each arm accounted for approximately 1,000 patient-days per month.
[f] Dosages of study drugs were selected based on patients' weight, which was visually estimated by the treating clinician.[15]
[g] The study followed enrolled patients over the first 6 days of their hospitalizations. Baseline was day 1, and the following 5 days (days 2-6) represented the follow-up period.[19]
[h] Patients' antipsychotic treatment was categorized as including any olanzapine or not, including any risperidone or not, and including any haloperidol or not. The three cohorts thus overlap, because each cohort included all patients who received the respective drug in any amount and at any time throughout the 5-day study period.[19]
[i] Patients may have received more than one diagnosis or experienced more than one behavioral disturbance.[19]
[j] Behavioral and emotional disorders with onset usually occurring in childhood and adolescence; behavioral syndromes associated with physiological disturbances and physical factors; neurotic, stress-related and somatoform disorders; and mental retardation.[19]
BPRS = Brief Psychiatric Rating Scale; BVC = Brøset Violence Checklist; CI = confidence interval; CRT = cluster randomized trial; G = group; i.m. = intramuscular; i.v. = intravenous; kg = kilogram; mg = milligram; mins = minutes; n or N = number; NOS = not otherwise specified; NR = not reported; NRCT = nonrandomized controlled trial; NS = not significant;

PTSD = post-traumatic stress disorder; RCT = randomized controlled trial; SD = standard deviation; SDAS = Social Dysfunction and Aggression Scale; U.S. = United States; vs. = versus.

References

1. Kontio R, Pitkanen A, Joffe G, et al.: eLearning course may shorten the duration of mechanical restraint among psychiatric inpatients: a cluster-randomized trial. Nord J Psychiatry 68:443-449, 2014.
2. Smoot SL, Gonzales JL: Cost-effective communication skills training for state hospital employees. Psychiatric Services 46:819-822, 1995.
3. Abderhalden C, Needham I, Dassen T, et al.: Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial. British Journal of Psychiatry 193:44-50, 2008.
4. van de Sande R, Nijman HL, Noorthoorn EO, et al.: Aggression and seclusion on acute psychiatric wards: effect of short-term risk assessment. British Journal of Psychiatry 199:473-478, 2011.
5. Putkonen A, Kuivalainen S, Louheranta O, et al.: Cluster-randomized controlled trial of reducing seclusion and restraint in secured care of men with schizophrenia. Psychiatric Services 64:850-855, 2013.
6. Nurenberg JR, Schleifer SJ, Shaffer TM, et al.: Animal-assisted therapy with chronic psychiatric inpatients: equine-assisted psychotherapy and aggressive behavior. Psychiatric Services 66:80-86, 2015.
7. Carlson JM, Holm MB: Effectiveness of occupational therapy for reducing restraint use in a psychiatric setting. American Journal of Occupational Therapy 47:885-889, 1993.
8. Bieniek SA, Ownby RL, Penalver A, et al.: A double-blind study of lorazepam versus the combination of haloperidol and lorazepam in managing agitation. Pharmacotherapy 18:57-62, 1998.
9. Dorevitch A, Katz N, Zemishlany Z, et al.: Intramuscular flunitrazepam versus intramuscular haloperidol in the emergency treatment of aggressive psychotic behavior. American Journal of Psychiatry 156:142-144, 1999.
10. Georgieva I, Mulder CL, Noorthoorn E: Reducing seclusion through involuntary medication: a randomized clinical trial. Psychiatry Research 205:48-53, 2013.
11. Isbister GK, Calver LA, Page CB, et al.: Randomized controlled trial of intramuscular droperidol versus midazolam for violence and acute behavioral disturbance: the DORM study. Annals of Emergency Medicine 56:392-401 e391, 2010.
12. Krakowski MI, Czobor P, Citrome L, et al.: Atypical antipsychotic agents in the treatment of violent patients with schizophrenia and schizoaffective disorder. Archives of General Psychiatry 63:622-629, 2006.
13. Krakowski MI, Czobor P, Nolan KA: Atypical antipsychotics, neurocognitive deficits, and aggression in schizophrenic patients. Journal of Clinical Psychopharmacology 28:485-493, 2008.
14. Michaud CJ, Thomas WL, McAllen KJ: Early pharmacological treatment of delirium may reduce physical restraint use: a retrospective study. Annals of Pharmacotherapy 48:328-334, 2014.
15. Richards JR, Derlet RW, Duncan DR: Chemical restraint for the agitated patient in the emergency department: lorazepam versus droperidol. Journal of Emergency Medicine 16:567-573, 1998.
16. Villari V, Rocca P, Fonzo V, et al.: Oral risperidone, olanzapine and quetiapine versus haloperidol in psychotic agitation. Progress in Neuro-Psychopharmacology and Biological Psychiatry 32:405-413, 2008.
17. Volavka J, Czobor P, Nolan K, et al.: Overt aggression and psychotic symptoms in patients with schizophrenia treated with clozapine, olanzapine, risperidone, or haloperidol. Journal of Clinical Psychopharmacology 24:225-228, 2004.

18. Volavka J, Czobor P, Sheitman B, et al.: Clozapine, olanzapine, risperidone, and haloperidol in the treatment of patients with chronic schizophrenia and schizoaffective disorder. American Journal of Psychiatry 159:255-262, 2002.
19. Wilhelm S, Schacht A, Wagner T: Use of antipsychotics and benzodiazepines in patients with psychiatric emergencies: results of an observational trial. BMC Psychiatry 8:61, 2008.

Online Supplement Tables for Risk of Bias Ratings

We provide our detailed risk of bias (ROB) ratings and the questions used to assign ratings below. ROB rating information for randomized controlled trials is presented in Online Supplement Tables 3 through 6, while ROB rating information for observational and nonrandomized controlled trials is shown in Online Supplement Tables 7 through 9.

Online Supplement Table 3. Risk of bias assessments for RCTs, part 1

Author, Year Trial Name (if applicable)	Type of Randomization	Eligibility criteria clearly described?	Method of randomization method appropriate?	Allocation concealment adequate?	Patients blind to treatment assignment	Outcome assessors blind to txmt assignment?	Care providers blind to txmt assignment?	Any variation from study protocol?	Groups recruited over same time period?
Abderhalden et al., 2008[1]	Cluster	Yes	Yes	Yes	Unclear	No	No	No	Yes
Bieniek et al., 1998[2]	Parallel	Yes	Yes	Yes	Yes	Yes	Yes	NR	Yes
Dorevitch et al., 2008[3]	Parallel	Yes	No	Unclear	Yes	Yes	Unclear	No	Yes
Georgieva et al., 2013[4]	Parallel	Yes	Unclear	Unclear	Unclear	Yes	No	No	Yes
Isbister et al., 2010[5]	Parallel	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Kontio et al., 2014[6] (also contains link to online study protocol) Kontio et al., 2011[7]	Cluster	Yes	Partially (coin toss, not stratified to control for "difficult to manage wards" – i.e., both were assigned intervention)	Yes	NA	Yes	No	Unclear	Yes (see Kontio et al., 2011)
Krakowski et al., 2006[8] Krakowski et al., 2008[9] Krakowski et al., 2009[10]	Parallel	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes

Online Supplement Table 3. Risk of bias assessments for RCTs, part 1 (continued)

Author, Year Trial Name (if applicable)	Type of Randomization	Eligibility criteria clearly described?	Method of randomization method appropriate?	Allocation concealment adequate?	Patients blind to treatment assignment	Outcome assessors blind to txmt assignment?	Care providers blind to txmt assignment?	Any variation from study protocol?	Groups recruited over same time period?
Nurenberg et al., 2015[11]	Parallel	Yes	Unclear	Unclear	No	Yes (violent and nonviolent incidents, seclusion and restraint use, OAS scores, psychiatric symptom scales) No (staff expectation of AAT benefit)	No	Unclear	Unclear
Putkonen et al., 2013[12]	Cluster	Yes	Unclear	No Data	No	Unclear	No	No Data	Yes
Richards et al., 1998[13]	Parallel	Yes	Unclear	Yes	NR	No	No	NR	Unclear
Smoot et al., 1997[14]	Cluster	Partially (criteria for unit selection NR)	No	NA	NA	NA	No	NR	Yes
van de Sande et al., 2011[15]	Cluster	Yes	Yes	Yes	Unclear	No	No	No	Yes
Volavka et al., 2004[16] Volavka et al., 2002[17] Czobor et al., 2002[18]	Parallel	Yes	Unclear	Unclear	Yes	Yes	Yes	No	No (see Rationale)

AAT = animal-assisted therapy; NA = not applicable; NR = not reported; OAS = Overt Aggression Scale; txmt = treatment

Online Supplement Table 4. Risk of bias assessments for RCTs, part 2

Author, Year	Baseline chx similar? If not similar, did design or analyses account for this?	Interventions adequately described?	Intervention fidelity adequate?	Cross-overs or contamination raising concern for bias?	KQ 1 Primary Outcomes: Valid and reliable measures consistently used for all participants?	KQ 1 Secondary Outcomes: Valid and reliable measures consistently used for all participants?	KQ 1: Benefits outcome data clearly reported without discrepancies?
Abderhalden et al., 2008[1]	No, and design/analyses did not account for differences	Yes	No	Unclear	Yes	Yes	Yes
Bieniek et al., 1998[2]	Yes, similar characteristics	Yes	Yes	No	Yes	No (non-validated VAS)	Yes
Dorevitch et al., 2008[3]	Unclear	Yes	Yes	No	Yes	NA	Yes
Georgieva et al., 2013[4]	Yes, similar characteristics	Yes	Unclear	Unclear	Yes	Yes	Yes
Isbister et al., 2010[5]	No, and design/analyses did not account for differences	Yes	Unclear	No	Yes	Yes	Yes
Kontio et al., 2014[6] (also contains link to study protocol) Kontio et al., 2011[7]	No for patients; Unclear for staff; No for both patients and staff Wards not stratified by function, diagnostic profile, average length of stay, or other parameters	Yes	Unclear	No	Yes	NA (secondary outcomes only reported for original 12 enrolled units, not for 10 remaining units in assessment of rates and duration of seclusion and restraint)	Partially (error in Kontio et al., 2014, Table 2's baseline min and max seclusion rates for control wards)
Krakowski et al., 2006[8] Krakowski et al., 2008[9] Krakowski et al., 2009[10]	Yes	Yes	Unclear	No	Yes	Yes	Yes
Nurenberg et al., 2015[11]	Partially (higher OAS-M aggression and life skills dysfunction in	Yes	Unclear	No	Yes (frequency of aggressive behavior)	Yes	Yes

	EAP vs. CAP group)

Online Supplement Table 4. Risk of bias assessments for RCTs, part 2 (continued)

Author, Year	Baseline chx similar? If not similar, did design or analyses account for this?	Interventions adequately described?	Intervention fidelity adequate?	Cross-overs or contamination raising concern for bias?	KQ 1 Primary Outcomes: Valid and reliable measures consistently used for all participants?	KQ 1 Secondary Outcomes: Valid and reliable measures consistently used for all participants?	KQ 1: Benefits outcome data clearly reported without discrepancies?
Nurenberg et al., 2015[11]	Partially (covariance analyses for life skills dysfunction, but not OAS-M aggression)				No (OAS-M verbal and physical aggression scores)		
Putkonen et al., 2013[12]	No data	Yes	No Data	No Data	Yes	NA	Yes
Richards et al., 1998[13]	Yes, similar characteristics	Yes	NR	No	Unclear (consistent use)	NA	Yes
Smoot et al., 1997[14]	Partially (demographic and clinical chx unaccounted for), and No (baseline levels of assaults on staff)	Yes	No	No	Yes	Yes	Yes
van de Sande et al., 2011[15]	No, but design/analyses accounted for differences	Yes	Unclear	Yes	Yes	Yes	Yes
Volavka et al., 2004[16] Volavka et al., 2002[17] Czobor et al., 2002[18]	Yes	Yes	Unclear	No	Yes	Yes	Yes

CAP = canine-assisted psychotherapy; chx = characteristics; EAP = equine-assisted psychotherapy; KQ = Key Question; NA = not applicable; NR = not reported; OAS-M = Overt Aggression Scale-Modified for Outpatient; VAS = visual analogue scale.

Online Supplement Table 5. Risk of bias assessments for RCTs, part 3

Author, Year	KQ 2 Harms: Valid and reliable measures consistently used for all participants?	KQ 2: Harms outcome data clearly reported without discrepancies?	Important outcomes pre-specified? If yes, reported?	Overall attrition?	Differential attrition?	Differential (≥15%) or overall high attrition (generally ≥20%) raising concern for bias?
Abderhalden et al., 2008[1]	NA	NA	Yes	0%	0%	No
Bieniek et al., 1998[2]	NA	NA	Yes	0%	0%	No
Dorevitch et al., 2008[3]	NA	NA	Yes	0%	0%	No
Georgieva et al., 2013[4]	NA	NA	Yes	0%	0%	No
Isbister et al., 2010[5]	Yes	Yes	Yes	13%	11.3% (droperidol vs. droperidol plus midazolam); 6.9% (droperidol vs. midazolam); 4.4% (droperidol vs. midazolam)	No
Kontio et al., 2014[6] (also contains link to study protocol) Kontio et al., 2011[7]	NA	NA	Yes, yes	Staff training attrition rates based on 12 originally enrolled units (see Kontio et al., 2011): Overall for all randomized staff: 39.9% Overall for staff completing baseline completers only: 43.1% Training completion for all randomized staff: 12.2%	Staff training attrition rates based on 12 originally enrolled units (see Kontio et al., 2011): 15.6% for all randomized staff 34.3% for all staff completing baseline surveys only Unclear in 10 final units (see Kontio et al., 2014)	Yes for staff training attrition in 12 original units; Unclear in 10 final units (see Kontio et al., 2014)

				Training completion for		

Online Supplement Table 5. Risk of bias assessments for RCTs, part 3 (continued)

Author, Year	KQ 2 Harms: Valid and reliable measures consistently used for all participants?	KQ 2: Harms outcome data clearly reported without discrepancies?	Important outcomes pre-specified? If yes, reported?	Overall attrition?	Differential attrition?	Differential (≥15%) or overall high attrition (generally ≥20%) raising concern for bias?
Kontio et al., 2014[6] (also contains link to study protocol) Kontio et al., 2011[7]				staff completing baseline surveys only: 6.5% Unclear in 10 final units (see Kontio et al., 2014)		
Krakowski et al., 2006[8] Krakowski et al., 2008[9] Krakowski et al., 2009[10]	Yes	Partially (limited information reported about ethnic group differences in harm outcomes)	Yes, yes	36.4%	5.4% to 14.7%	Yes
Nurenberg et al., 2015[11]	NA	NA	Yes, yes	7.8%	1.4% to 8.9%	No
Putkonen et al., 2013[12]	Yes	Unclear	Yes	No Data	No Data	No Data
Richards et al., 1998[13]	Partially (vital signs not consistently measured)	Yes	Partially (adverse events not related to vital signs)	8.2% (missing or incomplete data)	1.8% (missing or incomplete data)	No
Smoot et al., 1997[14]	Yes	Yes	Yes	9.7% (pre- and post-testing); 46% of experimental group did not complete training	3.3% (pre- and post-testing)	Yes
van de Sande et al., 2011[15]	NA	NA	Yes	0%	0%	No
Volavka et al., 2004[16]	Yes	Yes	Yes, yes	42%	1.8% to 15.4%	Yes

Volavka et al., 2002[17]						

Online Supplement Table 5. Risk of bias assessments for RCTs, part 3 (continued)

Author, Year	KQ 2 Harms: Valid and reliable measures consistently used for all participants?	KQ 2: Harms outcome data clearly reported without discrepancies?	Important outcomes pre-specified? If yes, reported?	Overall attrition?	Differential attrition?	Differential (≥15%) or overall high attrition (generally ≥20%) raising concern for bias?
Czobor et al., 2002[18] (harms data only)						

CAP = canine-assisted psychotherapy; EAP = equine-assisted psychotherapy; KQ = Key Question; NA = not applicable; SSP = environmentally enhanced social skills group psychotherapy; vs. = versus

Online Supplement Table 6. Risk of bias assessments for RCTs, part 4

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Abderhalden et al., 2008[1]	Unclear	NA	No	No	Medium	At baseline, rates of aggression were higher in intervention wards; unclear if interventions were implemented because of risk assessment. Because the unit of randomization was the hospital ward, raters were not blinded to treatment allocation across multiple psychiatric hospitals. There were fewer patients with schizophrenia in the preference group but all other characteristics were similar between groups. There was no reporting of attrition or intervention fidelity. Authors did not describe how wards from multiple hospitals were handled in analyses. No control for confounding.
Bieniek et al., 1998[2]	NA	NA	Yes	No	Low	No missing data, no attrition, and use of adequate randomization and blinding all strengths of the study. Small sample size (N=20) limited the study's statistical power to evaluate between-group differences, possibly explaining the nonsignificant group-by-time interaction for improvement in OAS scores, despite time to improvement clearly favoring haloperidol + lorazepam. Also unclear which timeframe was used for patient enrollment.
Dorevitch et al., 2008[3]	NA	NA	No	No	Medium	The study was very small (N=28). Unclear whether important sociodemographic variables differed between the two arms (no demographic or other clinical parameters were described); no control of potential confounders between two arms. The authors don't report on treatment fidelity or contamination, but it is unlikely to be a large concern given the study's small size. Authors don't provide info on attrition, but it seems unlikely to be a problem given the population and setting.

Online Supplement Table 6. Risk of bias assessments for RCTs, part 4 (continued)

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Georgieva et al., 2013[4]	NA	NA	No	No	High	The authors did not provide any details about randomization procedures, and clinicians on unit were not clearly blinded from patient assignments to the intervention arm. Data on the use of restrictive measures were extracted from the hospital database, but it is unclear who did the extracting and if s/he was blind to the randomization. Could not collect reliable data on the number of aggressive incidents in each arm. Authors didn't appear to take into account repeated measures, nor did they report results for only first admission, even though 21% of patients were repeat patients. Unclear whether confounders were controlled for; all presented results are unadjusted. Nearly three-quarters of the patients in Group 1 (first-choice involuntary medication) were also secluded, suggesting contamination.
Isbister et al., 2010[5]	Unclear	NA	Yes		Medium	There were potential confounding variables not addressed in the analysis (e.g., gender). The effect of additional sedation (when needed) in the ITT sample vs. the completers sample receiving only their randomized medication was not described. Unclear how the physical restraints required with medication administration affected outcomes of interest. Unclear how missing data were handled.
Kontio et al., 2014[6] (also contains link to study protocol) Kontio et al., 2011[7]	Yes (for seclusion and restraint data)	Unclear	No	Yes (see Rationale)	High	Randomization failed to allocate units in a balanced fashion, resulting in different case mixes among units in each group. This was especially problematic because both units for difficult-to-manage patients received the training intervention. Investigators unable to stratify their analyses of units by function, diagnostic profile, average length of stay, or other parameters to account for this.

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
						High overall and differential attrition of staff during the intervention phase (January to November 2009), with greater attrition in the control units'

Online Supplement Table 6. Risk of bias assessments for RCTs, part 4 (continued)

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Kontio et al., 2014[6] (also contains link to study protocol) Kontio et al., 2011[7]						staff (48% of those randomized and 62% of baseline completers). Kontio et al. (2011) (parent study) describes unexpected improvement in control group's attitudes toward seclusion possibly resulting from this differential attrition, which may have affected seclusion/restraint use outcomes. Other notable sources of bias include 1) inability to stratify units by potential modifying characteristics and 2) potential impact of seclusion and restraint-focused education-as-usual in some participating control units.
Krakowski et al., 2006[8] Krakowski et al., 2008[9] Krakowski et al., 2009[10]	Yes – ITT (see Krakowski et al., 2006, "Statistical Analyses")	Yes (see Rationale)	Yes	No	High	High overall attrition (36.4%) and differential attrition between haloperidol and olanzapine groups (14.7%), and ITT analyses likely not enough to offset the resulting bias. Unclear how ITT findings would differ from completers analysis. Also, potential bias from investigators' decision to pool first and second study sites likely minimal because second site only enrolled 7.3% of sample (8 patients). An otherwise well-designed study that took measures to minimize effects of important potential confounders and co-interventions. Strengths included, but were not limited to, double-blinded benefit and harm outcome assessment and medication administration for study drugs and benztropine for EPS.
Nurenberg et	Unclear	NA	Partially (see Rationale)	Yes (see	Medium	Randomized study with low overall and differential

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
al., 2015[11]				Rationale)		attrition that took multiple steps to reduce potential confounding from affecting results. Inclusion of active control group helped to control for potential impact of environmental changes on efficacy. Patients and intervention providers not blinded to group assignment, but not feasible given the types of interventions being provided. Issues potentially introducing bias include baseline OAS-M scores not adjusted for in covariance analyses

Online Supplement Table 6. Risk of bias assessments for RCTs, part 4 (continued)

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Nurenberg et al., 2015[11]						(significantly different between EAP and CAP groups), effect of patient expectations on intervention outcomes (not assessed), and inability to match patients in AAT groups with preferred animals.
Putkonen et al., 2013[12]	No Data	NA	Yes	No	Medium	To avoid unbalanced comparisons, intervention and control wards were stratified by use of seclusion and restraint. One senior psychiatrist, not associated with the study, made all pharmacological decisions in both wards. The unit of randomization was the hospital ward; as such, there was no blinding of treatment allocation. It was unclear, though probable, that the outcome examiners knew which ward the patient came from (or if the ward had been randomized to 6 Core Strategies) based on health records. There was minimal control of confounding.
Richards et al., 1998[13]	No	NA	Partially (ethanol intoxication evaluated as confounder, but not physician seniority)	Yes	High	Potentially biased assessment of sedation (depending on primary clinician's experience), neither outcome assessors nor care providers blinded to treatment assessment, potential impact of uncontrolled intracorrelation within subjects across timepoints because of choice of statistical

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
						test, and small potential increase in ROB because missing data were excluded without ITT analysis or ensuring no impact from those data.
Smoot et al., 1997[14]	NR	NA	No	Yes	High	Small CRT involving only two units "randomized" and within the units, only 72 employees. No information about randomization besides use of coin flip. Eligibility criteria not described. Baseline similarity of staff demographics, or patient demographics or clinical characteristics also not described. Approximately 10% of staff refused pre- and post-testing; this probably did not affect the primary outcome of interest. However, almost half of the experimental group failed to complete the

Online Supplement Table 6. Risk of bias assessments for RCTs, part 4 (continued)

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Smoot et al., 1997[14]						intervention training (46% overall, with 40% in day shift and 17% in evening shift), which could potentially be a source of bias. Differences in outcomes could have also varied by shift because of difference in noncompletion rates. No description of how missing data from pre- and post-testing were addressed, nor the extent to which pre- and post-testing non-completers in the experimental group completed the training.
van de Sande et al., 2011[15]	Unclear	NA	Yes	No	Medium	There was a risk of rater bias because same nurses who used Crisis Monitor scale as part of intervention also evaluated aggression and seclusion outcomes. The authors state potential risk of contamination, but they make a case that notification of control ward nurses by intervention ward nurses likely did not impact outcome. Analysis controlled for potentially confounding measures.
Volavka et al., 2004[16]	Yes – ITT(see Volavka et al.,	Yes (see Volavka et al., 2002,	Yes	Yes (see Rationale)	High	High overall attrition (42%) and differential attrition between risperidone and olanzapine groups

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Volavka et al., 2002[17] Czobor et al., 2002[18]	2002, Table 1 and "Measures of Efficacy")	"Measures of Efficacy")				(15.4%), and LOCF ITT analyses likely not enough to offset the resulting bias. Also, investigators' decision to limit analyses by omitting first 24 days of data likely not made a priori. Additionally, possible cohort effect following introduction of olanzapine arm 17 months after start of trial (olanzapine arm added in Nov 1997, when evaluation of other arms started in June 1996), but investigators did not find evidence of differences in PANSS scores before vs. after adding this last arm. Unclear if incidence and OAS Total Aggression Severity scores (our primary outcome of interest) were affected. Otherwise, well-designed study that took measures to minimize or eliminate effects of important potential confounders and co-interventions.

Online Supplement Table 6. Risk of bias assessments for RCTs, part 4 (continued)

Author, Year	Appropriate statistical method for missing data?	If multicenter study, accounted for in analysis?	Potential confounders and modifying variables taken into account in design and/or analysis?	Other potential sources of bias?	ROB	Rationale for ROB Rating
Volavka et al., 2004[16] Volavka et al., 2002[17] Czobor et al., 2002[18]						

AAT = animal-assisted therapy; CAP = canine-assisted psychotherapy; CRT = cluster randomized trial; EAP = equine-assisted psychotherapy; EPS = extrapyramidal symptoms; ITT = intent-to-treat; LOCF = last observation carried forward; N = number of patients; NA = not applicable; NR = not reported; OAS = Overt Aggression Scale; OAS-M = Overt Aggression Scale-Modified for Outpatient; PANSS = Positive and Negative Syndrome Scale; RCT = randomized controlled trial; ROB = risk of bias; vs. = versus

Online Supplement Table 7. Risk of bias assessments for observational studies and nonrandomized controlled trials, part 1

Author, Year Trial Name (if applicable)	Study Design	Eligibility criteria clearly described?	Eligibility criteria measured with valid and reliable measures, consistently across all participants?	Strategy for recruiting participants different across groups?	Sample size sufficient to detect meaningfully significant differences?	Interventions adequately described?	Important outcomes pre-specified? If yes, reported?	Comparison group selection appropriate?[a]	Any attempt to balance patient allocation between groups?	Impacts from concurrent interventions or unintended exposures that might bias results ruled out?
Carlson et al., 1993[19]	Cohort (retro-spective)	Yes	Yes	No, retrospective chart review	No Data	No	Yes	No	Unclear	No
Michaud et al., 2014[20]	Cohort (retro-spective)	Yes	Unclear	No	Yes	No	Yes	Yes	No	No
Villari et al., 2008[21]	NRCT	Unclear	Unclear	No	Unclear (powered to detect very small differences, not necessarily clinically meaningful differences)	Yes	Partially (harms not pre-specified)	Yes	Yes	Partially (unclear if other medications taken in addition to what patients were assigned)
Wilhelm et al., 2008[22]	NRCT	Yes	Partially (NR which clinical diagnoses deemed eligible)	No	Yes	Partially (antipsychotic dosing NR)	Yes	Yes	No	No

[a] After taking into account feasibility and ethical considerations.

NR = not reported; NRCT = nonrandomized controlled trial.

Online Supplement Table 8. Risk of bias assessments for observational studies and nonrandomized controlled trials, part 2

Author, Year Trial Name (if applicable)	Study Design	Outcome assessors blind to txmt or exposure status?	Interventions/exposures assessed using valid and reliable measures, consistently across all participants?	Follow-up length sufficient to support benefits/harms evaluation?	Overall attrition?	Differential attrition?	Differential (≥15%) or overall high attrition (generally ≥20%) raising concern for bias?
Carlson et al., 1993[19]	Cohort (retro-spective)	No	No (dependent on accuracy of chart documentation)	Yes (90 days)	NA (charts selected for study based solely on meeting eligibility criteria)	NA	NA
Michaud et al., 2014[20]	Cohort (retro-spective)	No	Unclear	Yes	0%	0%	No
Villari et al., 2008[21]	NRCT	Yes	Partially (dosing distribution NR)	Yes (benefits), and Partially (harms)	9.9%	0.8% to 2.8%	No
Wilhelm et al., 2008[22]	NRCT	NR	Unclear	Yes	2.9%	0.2% to 2.2%	No

NR = not reported; NRCT = nonrandomized controlled trial; txmt = treatment

Online Supplement Table 9. Risk of bias assessments for observational studies and nonrandomized controlled trials, part 3

Author, Year Trial Name (if applicable)	Study Design	Confounding and/or effect modifying variables assessed using valid and reliable measures, consistently across all participants?	KQ 1: Appropriate statistical methods used for assessing primary benefit outcomes?	KQ 2: Appropriate statistical methods used for assessing harms outcomes?	Any impt information about primary outcomes missing?	ROB	Rationale for ROB Rating

Author, Year Trial Name (if applicable)	Study Design	Confounding and/or effect modifying variables assessed using valid and reliable measures, consistently across all participants?	KQ 1: Appropriate statistical methods used for assessing primary benefit outcomes?	KQ 2: Appropriate statistical methods used for assessing harms outcomes?	Any impt information about primary outcomes missing?	ROB	Rationale for ROB Rating
Carlson et al., 1993[19]	Cohort (retro-spective)	No	Yes	NA	Yes (see Rationale)	High	High risk of misclassification bias due to inconsistent documentation of occupational therapy (OT) exposure. Intervention applied based on "screening or word of mouth". When identified in charts, frequency and intensity of sessions not consistently described. Minimal distinction between eligibility criteria for OT and no-OT groups. Also, no reporting of extent to which patients' histories of aggression affected outcomes. Readmitted patients functionally excluded from study because "charts were unavailable". Also, no evaluation of frequency of seclusion and restraint episodes.
Michaud et al., 2014[20]	Cohort (retrospective)	Unclear	Yes	Yes	No	High	Unclear how many patients did not receive screen. Medication dosing is unknown. There was no control for differences in concomitant medication use between arms; no control for confounding in primary analyses. Unclear if/how restraint assessment was consistently applied. Study was powered to detect a 20% difference in primary outcome. There were no major differences between groups except a much higher percentage of hypervigilance documented in the treatment group. More than half of patients with at least 1 positive delirium score were not enrolled (mostly due to lack of mechanical ventilation, some due to missing data), which has the potential to bias the results.
Villari et al., 2008[21]	NRCT	Partially (unclear to what extent medication dosing varied by treating physician)	Yes	Yes	No	Medium	Baseline characteristics similar despite alternating assignment to medication groups. Authors accounted for several potential confounders, such as prior depot

Online Supplement Table 9. Risk of bias assessments for observational studies and nonrandomized controlled trials, part 3 (continued)

Author, Year Trial Name (if applicable)	Study Design	Confounding and/or effect modifying variables assessed using valid and reliable measures, consistently across all participants?	KQ 1: Appropriate statistical methods used for assessing primary benefit outcomes?	KQ 2: Appropriate statistical methods used for assessing harms outcomes?	Any impt information about primary outcomes missing?	ROB	Rationale for ROB Rating
Villari et al., 2008[21]							antipsychotic or ECT treatment, but unable to account for other concurrent treatments that patients might have been taking at time of admission. Doses determined by treating physicians, who might have also introduced bias that way. Open-ended collection of KQ 2 harms data possibly affected by bias and might have led to underreporting. Also unclear if time of individual patients' enrollment introduced any ROB.
Wilhelm et al., 2008[22]	NRCT	Partially (benzodiazepine use measured, but unclear how other potential confounders measured)	Partially (exploratory analyses only, no adjustment for use of multiple centers)	Partially (exploratory analyses only, no adjustment for use of multiple centers)	Partially (differences in monotherapy groups NR)	High	High risk of selection bias into different medication groups based on clinical indication and variables related to the treating physician. Unadjusted confounding by concomitant and prior medication use. No attempts to adjust statistically for between-hospital differences or patients' baseline demographic or clinical chx.

chx = characteristics; ECT = electroconvulsive therapy; impt = important; KQ = Key Question; NR = not reported; NRCT = nonrandomized controlled trial; ROB = risk of bias

References

1. Abderhalden C, Needham I, Dassen T, et al.: Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial. British Journal of Psychiatry 193:44-50, 2008.
2. Bieniek SA, Ownby RL, Penalver A, et al.: A double-blind study of lorazepam versus the combination of haloperidol and lorazepam in managing agitation. Pharmacotherapy 18:57-62, 1998.
3. Dorevitch A, Katz N, Zemishlany Z, et al.: Intramuscular flunitrazepam versus intramuscular haloperidol in the emergency treatment of aggressive psychotic behavior. American Journal of Psychiatry 156:142-144, 1999.
4. Georgieva I, Mulder CL, Noorthoorn E: Reducing seclusion through involuntary medication: a randomized clinical trial. Psychiatry Research 205:48-53, 2013.
5. Isbister GK, Calver LA, Page CB, et al.: Randomized controlled trial of intramuscular droperidol versus midazolam for violence and acute behavioral disturbance: the DORM study. Annals of Emergency Medicine 56:392-401 e391, 2010.
6. Kontio R, Pitkanen A, Joffe G, et al.: eLearning course may shorten the duration of mechanical restraint among psychiatric inpatients: a cluster-randomized trial. Nord J Psychiatry 68:443-449, 2014.
7. Kontio R, Lahti M, Pitkanen A, et al.: Impact of eLearning course on nurses' professional competence in seclusion and restraint practices: a randomized controlled study (ISRCTN32869544). Journal of Psychiatric and Mental Health Nursing 18:813-821, 2011.
8. Krakowski MI, Czobor P, Citrome L, et al.: Atypical antipsychotic agents in the treatment of violent patients with schizophrenia and schizoaffective disorder. Archives of General Psychiatry 63:622-629, 2006.
9. Krakowski MI, Czobor P, Nolan KA: Atypical antipsychotics, neurocognitive deficits, and aggression in schizophrenic patients. Journal of Clinical Psychopharmacology 28:485-493, 2008.
10. Krakowski M, Czobor P, Citrome L: Weight gain, metabolic parameters, and the impact of race in aggressive inpatients randomized to double-blind clozapine, olanzapine or haloperidol. Schizophrenia Research 110:95-102, 2009.
11. Nurenberg JR, Schleifer SJ, Shaffer TM, et al.: Animal-assisted therapy with chronic psychiatric inpatients: equine-assisted psychotherapy and aggressive behavior. Psychiatric Services 66:80-86, 2015.
12. Putkonen A, Kuivalainen S, Louheranta O, et al.: Cluster-randomized controlled trial of reducing seclusion and restraint in secured care of men with schizophrenia. Psychiatric Services 64:850-855, 2013.
13. Richards JR, Derlet RW, Duncan DR: Chemical restraint for the agitated patient in the emergency department: lorazepam versus droperidol. Journal of Emergency Medicine 16:567-573, 1998.
14. Smoot SL, Gonzales JL: Cost-effective communication skills training for state hospital employees. Psychiatric Services 46:819-822, 1995.
15. van de Sande R, Nijman HL, Noorthoorn EO, et al.: Aggression and seclusion on acute psychiatric wards: effect of short-term risk assessment. British Journal of Psychiatry 199:473-478, 2011.
16. Volavka J, Czobor P, Nolan K, et al.: Overt aggression and psychotic symptoms in patients with schizophrenia treated with clozapine, olanzapine, risperidone, or haloperidol. Journal of Clinical Psychopharmacology 24:225-228, 2004.
17. Volavka J, Czobor P, Sheitman B, et al.: Clozapine, olanzapine, risperidone, and haloperidol in the treatment of patients with chronic schizophrenia and schizoaffective disorder. American Journal of Psychiatry 159:255-262, 2002.

18. Czobor P, Volavka J, Sheitman B, et al.: Antipsychotic-induced weight gain and therapeutic response: a differential association. Journal of Clinical Psychopharmacology 22:244-251, 2002.
19. Carlson JM, Holm MB: Effectiveness of occupational therapy for reducing restraint use in a psychiatric setting. American Journal of Occupational Therapy 47:885-889, 1993.
20. Michaud CJ, Thomas WL, McAllen KJ: Early pharmacological treatment of delirium may reduce physical restraint use: a retrospective study. Annals of Pharmacotherapy 48:328-334, 2014.
21. Villari V, Rocca P, Fonzo V, et al.: Oral risperidone, olanzapine and quetiapine versus haloperidol in psychotic agitation. Progress in Neuro-Psychopharmacology and Biological Psychiatry 32:405-413, 2008.
22. Wilhelm S, Schacht A, Wagner T: Use of antipsychotics and benzodiazepines in patients with psychiatric emergencies: results of an observational trial. BMC Psychiatry 8:61, 2008.