Background

Aggressive Behavior

Aggressive behavior connotes using actual physical violence toward self, others, or property or making specific imminent verbal threats.1 In health care settings, approaches for actively aggressive patients have historically involved using either seclusion (involuntary placement of a patient in a locked room or area from which the patient is not allowed to leave) or restraints (involuntary administration of mechanical, pharmacologic, or physical interventions, which is seen as more restrictive than seclusion); these practices continue today.2,3 Since the late 1990s, the U.S. Centers for Medicaid & Medicare Services (CMS3) and the Joint Commission (www.jointcommission.org4) have required using seclusion and restraints only for a behavior that “jeopardizes the immediate physical safety of the patient, a staff member, or others”5 (including other patients) and when less restrictive measures have failed. Despite practice guidelines advocating limitations of seclusion or restraints as much as possible,6 data in the United States and Europe show that 10 percent to 30 percent of patients (adolescents, adults, and elderly persons) admitted to acute psychiatric units receive these interventions.7,9

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Deciding to use seclusion or restraints raises several significant clinical or policy issues. First is how to best balance the benefits and risks of seclusion or restraints with those of various alternatives to those practices.7 Second, whether an evidence
base even exists to support using seclusion or restraints is debatable. Third, usual care, often represented in comparative studies as whatever was done before a new intervention was tried, varies substantially. Most guidelines and standards from regulatory agencies and accrediting bodies now recommend using seclusion and restraints only as a last resort. Finally, using seclusion and restraints is closely followed as a quality-of-care measure, particularly for psychiatric patients in hospital settings.

**Treatment Strategies**

Much interest now focuses on using alternatives to seclusion and restraints. These strategies can address preventing aggressive behavior or reducing aggressive behavior once it has already developed (or both). Most alternatives are strongly influenced by the National Association of State Mental Health Program Directors’ Six Core Strategies. These Six Core Strategies ultimately aim to forestall or at least decrease aggressive behavior.

*Preventing aggressive behavior.* Preventive strategies can be either general, multicomponent interventions that apply to all individuals (whether or not they are aggressive) or specific procedures aimed at persons who are at especially high risk of becoming aggressive. General preventive strategies emphasize providing a calm environment in which aggression is less likely to develop and tend to focus on entire care units. They include the following: risk assessment; milieu-based changes such as sensory rooms, which provide a calm and supportive environment for patients; staffing changes, such as increased staff-to-patient ratios; specific staff training programs; and peer-based interventions. Specific preventive strategies often try to intercede at the point of agitation, which is seen as a risk factor for becoming aggressive. These techniques can involve supportive (often referred to as nonconfrontational) language and other verbal de-escalation techniques, cognitive behavioral techniques, pharmacologic intervention treating the underlying psychiatric illness, and recognition of triggers for aggressive behavior. These two preventive approaches can overlap; specific strategies may also be applied as a general approach on a unit-wide basis.

*Managing acute aggression.* If patients do become actively aggressive, clinicians can use either seclusion or restraints or alternative strategies. In such cases, alternatives can include emergency response teams; these encompass behavioral emergency response teams, rapid response teams, and psychiatric emergency response teams. In addition, clinicians can employ pharmacologic interventions to reduce agitation quickly (rather than more gradually treating the underlying illness).

**Scope and Key Questions**

**Scope of the Review**

This small systematic review addresses interventions to prevent or de-escalate aggressive behavior and to reduce use of seclusion and restraint for aggressive behaviors.

We focus on studies in acute health care settings, as to our knowledge no such review has been done using data from such settings. We are concerned with (1) the effectiveness of different available alternative strategies to prevent aggressive behavior and with (2) the effectiveness of alternative strategies compared with each other or with seclusion and restraints to de-escalate aggressive behaviors or improve health outcomes for those who are acutely aggressive. We conceptualize “de-escalate” in terms of both preventing aggressive behaviors and reducing use of seclusion and restraints.

We do not assess the accuracy of available risk assessment tools (a crucial step in the process of reducing aggressive behavior) or consider chronic care settings; although these are important considerations, they are beyond the scope of this review.

**Key Questions**

For the three Key Questions (KQs) in this review, we define aggressive behavior as making specific imminent verbal threats or using actual physical violence toward self, others, or property. We focus on patients with any psychiatric diagnosis per the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition-Revised, Fourth Edition, or Fifth Edition* (DSM-III-R, DSM-IV, or DSM-5). Diagnostic categories include delirium and substance misuse but not dementia; additionally, for patients in emergency departments, we include displaying severe psychiatric symptomatology. We view effectiveness in terms of both benefits and harms, so we frame our questions to address each class of outcomes.

We envision a continuum of risk and behavior, so the KQs cover a range of patients. This spectrum can include patients with these disorders who may be at risk of aggressive behavior (i.e., not actively aggressive), in which case interventions are preventive. It can also include those who are exhibiting aggressive behaviors (i.e., are actively aggressive), in which case interventions are directly active. Interventions can occur at any point along this continuum, and they can involve a wide variety of strategies that can have educational, behavioral, emotional, organizational, environmental, and/or pharmacologic components. The interventions must target a reduction either in aggressive behavior or in
use of seclusion and restraints. For these KQs, we define and classify interventions to reflect either prevention or direct intervention. A preventive intervention is one applied to a group of individuals not wholly identified as being actively aggressive; i.e., some patients may not be actively aggressive but others may be. It can involve unit- or hospital-wide policies that address all patients on a unit, not just those who are actively aggressive. It can also involve patients identified as being at an increased risk of becoming aggressive (e.g., were assessed as being agitated) but who were not yet actively aggressive.

KQ 1 (benefits) and KQ 2 (harms) address such preventive interventions in these groups in their subquestion (a). KQs 1 and 2, in their subquestions (b) and (c), examine interventions targeted specifically to de-escalate aggressive behavior among actively aggressive patients. KQ 3 addresses specific patient, intervention, or setting factors that may modify benefits or harms of various strategies.

Our two primary comparative outcome benefits (KQ 1), which are intermediate outcomes, are a decrease in (1) aggressive behaviors and (2) use of seclusion and restraints. We also look at longer term or final health outcomes. These include improved quality of life, functioning, or patient experience; improved therapeutic relationship; decreased subsequent aggressive behavior; and decreased subsequent use of seclusion and restraints. We also consider general resource use.

Acute health care settings are defined as public and private mental hospitals, acute care units at state mental hospitals, acute care components of Veterans Health Administration (VA) hospitals, medical or surgical units in general hospitals, and emergency departments. In all cases, patient discharges occur within 35 days of beginning treatment. Stays longer than 35 days would indicate a chronic care setting.

The three KQs are stated fully just below. Figure A then presents our analytic framework that guided this review; it identifies specific KQs.

KQ 1: Regarding benefits for adult psychiatric patients in acute care settings:

a. For those without active aggression, what are the comparative benefits of strategies to prevent aggressive behavior?

b. For those with active aggression, what are the comparative benefits of strategies, including seclusion and restraints, to de-escalate aggressive behavior?

c. For those with active aggression, what are the comparative benefits of strategies to reduce the use of seclusion and restraints?

KQ 2: Regarding harms for adult psychiatric patients in acute care settings:

a. For those without active aggression, what are the comparative harms of strategies to prevent aggressive behavior?

b. For those with active aggression, what are the comparative harms of strategies, including seclusion and restraints, to de-escalate aggressive behavior?

c. For those with active aggression, what are the comparative harms of strategies to reduce the use of seclusion and restraints?

KQ 3: What characteristics of patients (including age, sex or gender, diagnosis, motivation to receive treatment), of intervention components, or of acute care settings modify the benefits or harms of interventions for psychiatric patients at risk of, or presenting with, active aggression?
Analytic Framework

Figure A. Analytic framework for comparative effectiveness of strategies to de-escalate aggressive behavior in psychiatric patients

Methods

Topic Refinement and Protocol Review

During topic refinement we developed a draft and then a final review protocol. Specifically, we generated an analytic framework, preliminary KQs, and preliminary inclusion/exclusion criteria; these reflect PICOTS constructs (patients or populations, interventions, comparators, outcomes, time frames, and settings) and other details about eligible studies. Information from the topic nominator helped guide our processes. A panel of 10 Key Informants (KIs) gave input on the scope and details of initial KQs; these KQs were posted on the Agency for Healthcare Quality and Research (AHRQ) Web site for public comment (www.effectivehealthcare.ahrq.gov) from June 8, 2015, through June 29, 2015. We then revised the KQs as needed.

In addition, we consulted with seven experts (members of a Technical Expert Panel), who provided feedback as we developed our review protocol. Their inputs addressed points such as sample size thresholds for eligible studies and whether and how to limit assessments of risk of bias of individual studies.

Literature Search Strategy

Search Strategy and Eligibility Criteria

To identify relevant KQ-specific articles, we searched MEDLINE® (via PubMed), Embase®, the Cochrane Library, Academic Search Premier, PsycINFO, and CINAHL (Cumulative Index to Nursing and Allied Health Literature) from January 1, 1991, through February 3, 2016. Appendix A (main report) presents the full search strategy (limiting searches to English and human-only...
studies). An experienced information scientist—our Evidence-based Practice Center (EPC) librarian—ran all searches.

Our searches focused on comparative studies of de-escalation strategies (seclusion, restraints, or alternatives to seclusion or restraints) for patients with psychiatric disorders or severe psychiatric symptomatology who are at risk of, or presenting with, aggressive behavior across various acute care settings. Search strings included various Medical Subject Heading (MeSH) terms for psychiatric disorders, acute care settings, and aggressive behavior. We also manually searched reference lists of pertinent reviews, included trials, and background articles to identify relevant citations that our searches might have missed. To find relevant gray literature we followed guidance from the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews for these steps.

We developed inclusion/exclusion criteria with respect to PICOTS and study designs. Inclusion criteria limited populations to patients 18 years of age or older; they included any psychiatric or substance use disorder and delirium. Studies limiting populations to patients with dementia were ineligible.

We required that interventions target reducing aggressive behavior or decreasing use of seclusion and restraints (or both). Eligible studies had to have reported on at least one of our two primary outcomes: (1) decreased aggression in terms of frequency, severity, or duration (measured by either direct counts or validated aggression scales) or (2) reduced use of seclusion or restraints (decreased rate, amount, or duration). Investigators had to have tested interventions in acute care settings (general hospitals, psychiatric hospitals, and emergency departments in these hospitals).

**Study Selection**

Two members of the research team independently reviewed all titles and abstracts (generated by searches) against our inclusion/exclusion criteria. For evaluating the full text of publications, we retrieved those that either reviewer marked for inclusion and those without adequate information in titles or abstracts. Then, two investigators independently reviewed the full text to determine final inclusion or exclusion. The reviewers resolved any disagreements by discussion and consensus or by consulting a third member of the review team.

**Data Extraction**

We abstracted the following data from included trials and studies: study designs, eligibility criteria, population characteristics (such as age, sex, race, ethnicity), interventions, comparators, additional medications or interventions allowed, outcomes of interest and methods of outcome assessment, sample sizes, attrition, settings, geographic locations, and study funders. We recorded intention-to-treat results (i.e., all patients were analyzed as randomized with missing values imputed) if available. We resolved discrepancies by consensus or by involving a third, senior reviewer. When eligible studies reported data that were incomplete or missing, we contacted authors.

**Risk of Bias Assessment**

To assess the risk of bias of trials and certain other studies, we followed EPC methods guidance and rated the risk of bias for each relevant outcome as low, medium, or high. To determine risk of bias in a standardized way, we used the Cochrane Risk of Bias tool to appraise randomized controlled trials (RCTs). We also used it to appraise the few cluster randomized trials (hereafter CRTs, where clusters were based on specific units in the facilities where the studies took place). Guidance for assessing risk of bias is similar for RCTS and CRTs but the latter may need special attention to issues such as recruitment bias, baseline imbalance, loss of clusters, and inadequate or incorrect analytic techniques, and we made an effort to consider these matters in reviewing eligible CRTs. For nonrandomized trials and observational studies, we employed criteria from the RTI Risk of Bias Tool for Observational Studies. To minimize risk of bias in observational and noncontrolled studies addressing adverse outcomes (i.e., harms, a key focus of our report), we required a minimum total sample of 100 patients in nonrandomized studies (consistent with our work in prior reviews). We did not assess risk of bias in noncontrolled or pre/post studies.

Two independent reviewers assigned risk of bias ratings. Disagreements were resolved by discussion and consensus or by consulting a third, senior reviewer.

**Data Synthesis**

We synthesized all literature qualitatively, and included all eligible studies regardless of risk of bias. We stratified study data by whether they came from controlled studies (e.g., RCT, cohort studies) or noncontrolled studies (e.g., pre/post, interrupted time series).
A study might report data relevant to both preventive measures (subquestion [a]) and actively aggressive measures (subquestion [b] or [c]). Data for study groups not restricted to highly aggressive patients (i.e., the denominator involved both aggressive and nonaggressive patients) were considered relevant for subquestion (a). Data for groups restricted to highly aggressive patients were considered relevant to subquestions (b) and (c).

To determine whether quantitative analyses (i.e., meta-analysis) were appropriate, we assessed the clinical and methodological heterogeneity of the studies under consideration following established guidance. After qualitatively assessing the PICOTS of included studies looking for similarities and differences, we determined that our body of evidence was too heterogeneous to justify quantitative analyses.

**Strength of the Body of Evidence**

We graded the strength of evidence (SOE) for primary outcomes based on the guidance established by the EPC Program. Developed to grade the overall strength of a body of evidence, this approach incorporates five key domains: study limitations (study design and aggregate risk of bias), consistency, directness, precision, and reporting bias. For some scenarios, this approach also considers other optional domains that may be relevant: a dose-response association, plausible confounding that would decrease the observed effect, and strength of association (magnitude of effect). SOE receives one of four grades: high, moderate, low, or insufficient. In grading evidence from single trials or studies (typically regarded as insufficient evidence), we gave more weight to those in which the reported findings were precise and graded some as low SOE. Mirroring our decision not to assess the risk of bias of pre/post studies, we did not grade the SOE from such studies, as they cannot be used to draw causal inferences about comparative benefits and harms. Two trained reviewers assessed each domain for each primary outcome; differences were resolved by consensus. One of the two reviewers was always a senior researcher with experience in grading SOE.

**Applicability**

We assessed applicability of the evidence following guidance from the Methods Guide for Effectiveness and Comparative Effectiveness Reviews. We used the PICOTS framework to explore factors that affect applicability. Some factors identified a priori that may limit the applicability of evidence include the following: age of enrolled populations, sex of enrolled populations (e.g., fewer men may be enrolled in some studies), race or ethnicity of enrolled populations, diagnoses of involved sample, and location of and staffing for specific interventions.

**Peer Review and Public Commentary**

This report was posted for public comment and peer review. We addressed all comments in the final report, making revisions as needed. A disposition of comments report will be publicly posted 3 months after release of the final report.

**Results**

**Literature Searches and Evidence Base**

 Searches of all sources identified a total of 1,921 potentially relevant citations. We included 29 primary studies (described in 31 articles) that compared interventions to de-escalate aggressive behavior or reduce use of seclusion or restraints with an alternative strategy or usual care and provided data for one or more KQs. Of these, 11 were controlled trials that provided eligible data for SOE ratings. Only 4 took place in the United States: 1 in an inpatient psychiatric unit, 2 in an emergency department, and 1 in an intensive care unit with intubated patients. The remaining 18 studies were pre/post studies, for which we did not grade SOE; we identified no interrupted time-series studies. We report below only on findings from trials or studies for which we could grade SOE.

We had data for KQs 1 (benefits) and 2 (harms) from the following types of trials or studies: KQ 1a (benefits of prevention), three CRTs; KQ 1b (benefits of de-escalating aggression), four RCTs and two nonrandomized controlled trials (NRCTs); KQ 1c (benefits of reducing seclusion/restraint use), one CRT and one retrospective cohort study; KQ 2a (harms of prevention), one CRT; and KQ 2b (harms of de-escalating aggression), four RCTs and two NRCTs. No eligible studies pertained to KQ 2c. We identified no eligible studies for KQ 3.

Most interventions took place in public psychiatric hospitals. For studies reporting on demographics for their patient populations, the mean age ranged primarily between 38 and 40 years, the distribution of men and women varied widely across studies, and race or ethnicity was sparsely reported.

We analyzed five broad categories of interventions: staff training; risk assessment; multimodal; environmental or group psychotherapeutic; and medication protocols. Studies that did not differentiate their results between...
those patients with aggression and those who were not yet aggressive were included in prevention analyses.

We found the SOE for most of the findings to be insufficient, with the justification for these assessments provided in the tables below (see Appendix D of the main report for detail about scores for each SOE domain). To help clarify this literature's range of different types of studies, and the heterogeneity of approaches, populations, settings, and outcomes, we report below the findings for all 11 eligible studies, whether the SOE was insufficient or low. We report the findings as the authors reported them; we then indicate the SOE for the finding.

**Comparative Benefits of Strategies**

**Key Question 1a: Benefits of Strategies to Prevent Aggressive Behavior**

**Staff Training Interventions Versus Usual Care**

Staff training in interpersonal communication led to fewer incidents of seclusion and restraint and a larger decrease in incidents of seclusion and restraint than usual care on a control unit\(^69\) (one CRT, insufficient SOE).

**Risk Assessment Interventions Versus Usual Care**

Units employing structured risk assessment protocols reported significantly fewer aggressive incidents than usual care units. One CRT focused on lowering severe aggressive incidents\(^44\), the other focused on any aggressive incidents\(^53\) (one CRT for each outcome, low SOE).

Cluster trials in which units employed structured risk assessment protocols reported significantly fewer hours spent in seclusion\(^53\) (one CRT, low SOE) and significantly fewer coercive measures than usual-care units\(^44\) (one CRT, low SOE).

**Multimodal Interventions Versus Usual Care**

No studies assessed multimodal interventions to prevent aggression in patients without active aggression (insufficient SOE).

**Environmental or Group Psychotherapeutic Interventions Versus Usual Care**

No studies assessed environmental or group psychotherapeutic interventions in patients without active aggression (insufficient SOE).

**Medication Protocols Versus Other Medication Protocols or Alternative Strategies**

No studies assessed medication protocols in patients without active aggression (insufficient SOE).

In Table A for KQ 1a, we present the supporting judgment for our SOE grades for evidence from studies with eligible study designs (i.e., any study that we could rate for risk of bias). Supporting judgment is essentially the ratings on the main domains for grading SOE (study limitations [i.e., risk of bias], consistency, directness, and precision). The CRTs in this report did not control for clustering in their statistical analyses, which weakened the SOE grade. Table A has entries only for staff training (one CRT) and for risk assessment strategies (two studies); we had no relevant studies for the other three types of interventions.
<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff training vs. usual care</td>
<td>Change in aggressive behavior</td>
<td>Aggressive behavior resulting in staff injury</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Fewer assaults on staff occurred in unit that received the staff training vs. the control unit (4 vs. 5); no statistical testing reported. ^6^</td>
</tr>
<tr>
<td>Change in seclusion or restraint</td>
<td>Incidents of seclusion or restraint</td>
<td>NR</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Fewer incidents of seclusion or restraint on the unit who received the training vs. the control unit (84 vs. 228), no statistical testing reported. ^6^</td>
</tr>
<tr>
<td>Risk assessment vs. usual care</td>
<td>Change in aggressive behavior</td>
<td>Number of aggressive patients: 170 during baseline period, 458 during intervention period</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Nonsignificant 50% RR reduction with risk assessment vs. usual care. ^5^</td>
</tr>
<tr>
<td></td>
<td>Aggressive incidents</td>
<td>170 during baseline period, 458 during intervention period</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant 68% RR reduction with risk assessment vs. usual care, p&lt;0.0001 reported; failure to control for intraclass correlations weakens the finding. ^5^</td>
</tr>
<tr>
<td></td>
<td>Rate of severe aggressive incidents</td>
<td>973 post-intervention</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significantly lower risk with structured risk assessment: (RR, 0.59; 95% CI, 0.41 to 0.83); p&lt;0.001 reported; failure to control for intraclass correlations weakens the finding. Decrease achieved since baseline with risk assessment (-41%) vs. usual care (-15%), no statistical testing reported. ^4^</td>
</tr>
<tr>
<td></td>
<td>Change in physical attacks</td>
<td>973 post-intervention</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significantly greater decrease with risk assessment (-41%) vs. usual care (7%), p&lt;0.001 reported, failure to control for intraclass correlations weakens the finding. ^4^</td>
</tr>
<tr>
<td>Secluded patients</td>
<td>170 during baseline period, 458 during intervention period</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Nonsignificant 8% RR increase with risk assessment vs. usual care. ^5^</td>
<td></td>
</tr>
</tbody>
</table>
### Table A. Summary of findings with strength of evidence grades: Comparative benefits of two strategies for preventing aggressive behaviora (KQ 1a) (continued)

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome</th>
<th>N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk assessment vs. usual care</td>
<td>Change in aggressive behavior</td>
<td>Seclusion incidents</td>
<td>170 during baseline period, 458 during intervention period</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Nonsignificant 15% RR reduction with risk assessment vs. usual care.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hours in seclusion</td>
<td>170 during baseline period, 458 during intervention period</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant 45% RR reduction with risk assessment vs. usual care, p&lt;0.0001 reported; failure to control for intraclass correlations weakens the finding.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change in coercivea incidents</td>
<td>973 post-intervention</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, indirect, precise</td>
<td>Significant decrease from baseline with risk assessment (-27%) vs. usual care (+10%), p&lt;0.001; failure to control for intraclass correlations weakens the finding.44</td>
</tr>
</tbody>
</table>

a For KQ 1a, we had no studies of eligible study design for environmental or group psychotherapeutic interventions or multimodal interventions; thus, we could not rate risk of bias.

b Coercive measures covered a wide range of measures, from forced injection of psychotropic medication to seclusion and mechanical restraint.44

CI = confidence interval; N = number; NR = not reported; RR = relative risk; vs. = versus.

### Key Question 1b: Benefits of Strategies to De-escalate Aggressive Behavior

No eligible studies targeted de-escalation using staff training, risk assessment, multimodal or environmental protocols. Thus, the SOE grades are insufficient for all such interventions.

Six studies assessed different medication protocols; the strength of evidence for each was insufficient. Four studies were RCTs. In an inpatient psychiatric unit, one RCT found no difference between the effects of intramuscular haloperidol versus intramuscular flunitrazepam for treating patients displaying aggressive psychotic behavior (one RCT, insufficient SOE). The remaining RCTs were in emergency department settings. One RCT in a public psychiatric hospital emergency department found that intramuscular droperidol for treating patients exhibiting violent and acute behavioral disturbance did not reduce the duration of aggressive behavior any more than intramuscular midazolam (one RCT, insufficient SOE). Another RCT in a hospital psychiatric emergency department found that, compared with intramuscular lorazepam, intramuscular lorazepam plus haloperidol for treating patients exhibiting serious, acute agitated, or aggressive behavior did not result in greater overall reduction of aggressive or agitated behavior, but the medication regimen did produce a more rapid reduction in aggressive or agitated behavior and more patients who achieved clinically significant improvement in aggressive or agitated behavior (one RCT, insufficient SOE). Finally, an RCT in an urban university emergency department found that intramuscular droperidol for intoxicated or psychiatrically ill, violently agitated patients requiring chemical restraint produced more rapid sedation and greater sedation overall than intramuscular lorazepam (one RCT, insufficient SOE).

Two studies were NRCTs. In an inpatient psychiatric hospital setting, treatments that included any olanzapine, any risperidone, or any haloperidol for treating patients with agitation did not differ from each other in reducing aggressive behavior or suicidality (one NRCT, insufficient SOE).
insufficient SOE). In an inpatient psychiatric emergency setting, the effects of oral risperidone, olanzapine, quetiapine, or haloperidol did not differ in reducing aggressive behavior (one NRCT, insufficient SOE).

In Table B for KQ 1b, we present information (supporting judgment) for our SOE grades for evidence based on studies with an eligible study design. For this subquestion, we had no relevant studies of staff training, risk assessment, multimodal, or environmental or group psychotherapeutic interventions. All findings for the medications protocols were underpowered to test noninferiority.

### Table B. Summary of findings with strength of evidence grades: Comparative benefits of medication protocols for de-escalating aggressive behavior (KQ 1b)

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome</th>
<th>N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication protocols vs. other medication protocols: Benefits</td>
<td>Change in aggressive behavior</td>
<td>Aggression response rate</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Nonsignificant difference in rates of OAS score reduction at 90 minutes in haloperidol vs. flunitrazepam (92% vs. 80%).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Duration of aggression</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Nonsignificant difference in the median duration of violent and acute behavioral disturbances with droperidol vs. midazolam vs. a combination of droperidol plus midazolam (20 vs. 24 vs. 25 minutes).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clinically significant change in OAS scores</td>
<td>Insufficient</td>
<td>Low risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Significantly greater likelihood of improvement (decrease of four or more points) in OAS scores of aggressive or agitated behavior at 60 minutes with the combination of haloperidol plus lorazepam (100%) vs. lorazepam alone (55%), p=0.03 (note small sample size).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time to OAS improvement</td>
<td>Insufficient</td>
<td>Low risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Significantly shorter time to OAS improvement with the combination of haloperidol plus lorazepam vs. lorazepam alone, data NR, p=0.028 (note small sample size).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table B. Summary of findings with strength of evidence grades: Comparative benefits of medication protocols for de-escalating aggressive behavior (KQ 1b) (continued)

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome</th>
<th>N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication protocols vs. other medication protocols: Benefits (continued)</td>
<td>Change in aggressive behavior (continued)</td>
<td>Sedation score at 5, 10, 15, 30, and 60 minutes</td>
<td>202</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significantly lower mean sedation scores (i.e., less combative, violent, or out of control behavior) at 10, 15, 30, and 60 minutes with droperidol vs. lorazepam, each p&lt;0.001.68</td>
</tr>
<tr>
<td>Change in CGI-A scores</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No differences in changes in percentages of patients with CGI-A score ≥3 from baseline to day 6 or to last day of observation with olanzapine vs. risperidone vs. haloperidol, p=NR.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in MOAS total aggression scores</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Nonsignificant differences between risperidone vs. olanzapine vs. quetiapine vs. haloperidol in changes in mean total MOAS scores from baseline to 72 hours.65</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CGI-A = Clinical Global Impression Severity of Illness – Aggression; MOAS = Modified Overt Aggression Scale; n = number of patients; NR = not reported; OAS = Overt Aggression Scale; vs. = versus.

### Key Question 1c: Benefits of Strategies to Reduce Seclusion and Restraint Use

No eligible studies addressed reductions in seclusion or restraints for staff training, risk assessment, multimodal, or environmental protocols. SOE grades were thus all insufficient.

Two high risk of bias studies addressed the comparative effectiveness of two different medication protocols to reduce seclusion and restraint use. In one study, in an inpatient psychiatric unit with acutely agitated and violent inpatients, a first choice of involuntary medication treatment with oral or intramuscular haloperidol plus promethazine was compared with a first choice of seclusion. The medication option did not produce differences either in subsequent mechanical restraint use52 (one RCT, insufficient SOE) or in subsequent coercive incidents (i.e., seclusion, restraint, or involuntary medications)52 (one RCT, insufficient SOE).

In the other study, for treating delirium in an inpatient intensive care unit, immediate (within 24 hours) treatment with at least one dose of an antipsychotic medication led to fewer mean days in restraints than did delayed or no treatment51 (one retrospective cohort, insufficient SOE).

In Table C on KQ 1c, we present the supporting judgment for our SOE grades for each eligible study (in this case only for medication protocols).
**Table C. Summary of findings with strength of evidence grades: Comparative benefits of medication-based strategies for reducing seclusion and restraint use in aggressive patients (KQ 1c)**

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication protocols vs. other medication protocols or usual care</td>
<td>Change in seclusion or restraint</td>
<td>Seclusion incident rate 659</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant lower risk with involuntary medication(^a) as first choice vs. seclusion as first choice (RR, 0.51; 95% CI, 0.34 to 0.79), p&lt;0.001(^52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seclusion hours 659</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Lower number of overall hours with involuntary medication(^a) as first choice vs. seclusion as first choice (998 vs. 2,098), no statistical testing reported(^52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seclusion duration 659</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Longer mean duration with involuntary medication(^a) as first choice vs. seclusion as first choice (32 vs. 30 hours), no statistical testing reported(^52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seclusion duration rate 659</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant lower risk with involuntary medication(^a) as first choice vs. seclusion as first choice (RR, 0.54; 95% CI, 0.5 to 0.58) p&lt;0.001(^52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mechanical restraint incident rate 659</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No significant difference in involuntary medication(^a) as first choice vs. seclusion as first choice (RR, 1.44; 95% CI, 0.38 to 5.36).(^52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coercive incident rate(^b) 659</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No significant difference in coercive incident rate in involuntary medication(^a) vs. seclusion as first choice options (RR, 0.95; 95% CI, 0.67 to 1.35).(^52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duration in restraints 200</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant decrease with single-dose delirium treatment vs. no delirium treatment, both in the first 24 hours, 3 vs. 6 days, p&lt;0.001(^51)</td>
</tr>
</tbody>
</table>

\(^a\) “Involuntary medication” refers to single dose haloperidol plus promethazine or lorazepam.

\(^b\) “Coercion” refers to a sequence of coercive episodes (seclusion, mechanical restraint, or involuntary medication) for less than 24 hours.

CI = confidence interval; KQ = Key Question; N = number; RR = relative risk; vs. = versus.
Key Question 2a: Harms of Strategies To Prevent Aggressive Behavior

No eligible studies examined risk assessments, multimodal interventions, environmental interventions, or medication protocols. SOE grades for these were insufficient.

One study addressed staff training. A unit on which staff received interpersonal communication training had fewer patient rights complaints, staff resignations and transfers, and sick leave than a control unit. Further, the intervention unit experienced a greater decrease in these outcomes during the study period than the control unit\(^69\) (one CRT, insufficient SOE).

Table D. Summary of findings with strength of evidence grades: Comparative benefits and harms of two strategies for preventing aggressive behavior (KQ 2a)

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome of Interest</th>
<th>N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff training vs. usual care</td>
<td>Staff distress</td>
<td>Change in staff distress</td>
<td>NR</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, indirect, imprecise</td>
<td>Fewer staff resignations and transfers in unit that received the staff training than in control unit (4 vs. 9), no statistical testing reported.(^69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change in staff sick leave</td>
<td>NR</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, indirect, imprecise</td>
<td>Greater percentage decrease in number of sick leave hours in unit that received the staff training than in control unit (28.2% vs. +7.7%), no statistical testing reported.(^69)</td>
</tr>
<tr>
<td>Patient distress</td>
<td>Change in patients’ rights complaints</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, indirect, imprecise</td>
<td>Fewer patients’ rights complaints occurred in unit that received the staff training than in control unit (2 vs. 4), no statistical testing reported.(^69)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = number; NR = not reported; vs. = versus.

Key Question 2b: Harms of Strategies To De-escalate Aggressive Behavior

No eligible studies tested staff training, risk assessments, multimodal, or environmental protocols.

Four RCTs and two NRCTs provided harms data for medication protocols; all reported small numbers of events and performed no statistical testing. These studies generally reported their harms findings as indicating no differences, but their studies were underpowered to test noninferiority. One RCT\(^43\) examined three possible harms: drug-related adverse effects; incidence of abnormal QT (QRS complex to T wave interval) interval; and incidence of staff injury after use of midazolam, droperidol, or their combination for patients with active aggression (one RCT, insufficient SOE). Another RCT reported on acute extrapyramidal events and incidence of marked sedation in a comparison between haloperidol and flunitrazepam\(^48\) (one RCT, insufficient SOE). A third RCT reported the incidence of side effects of lorazepam alone or in combination with haloperidol for adults treated in a psychiatric emergency service setting\(^64\) (one RCT, insufficient SOE). Finally, one RCT reported the incidence of differences in changes in vital signs in acutely agitated emergency department patients treated with droperidol or lorazepam\(^68\) (one RCT, insufficient SOE).

One NRCT reported the incidence of abnormal gait, dizziness, extrapyramidal events, headache, hypotension, or somnolence in 101 adult inpatients with psychosis receiving either risperidone, olanzapine, quetiapine, or haloperidol\(^65\) (one NRCT, insufficient SOE). Another
NRCT reported the incidence of treatment-emergent side effects, including extrapyramidal events, for patients receiving olanzapine, risperidone, or haloperidol\textsuperscript{72} (one NRCT, insufficient SOE).

Table E documents our SOE grades.

### Table E. Summary of findings with strength of evidence: Comparative harms of medication protocols for addressing aggressive behavior (Key Question 2b)

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication protocols vs. other medication protocols</td>
<td>Staff harm</td>
<td>Staff injury</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Very small numbers of events with no statistical testing for detecting differences in droperidol vs. midazolam vs. a combination of droperidol plus midazolam (3 vs. 1 vs. 2, p=NR).\textsuperscript{43}</td>
</tr>
<tr>
<td>Adverse effects from medication</td>
<td>Acute extrapyramidal events</td>
<td>28</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No acute extrapyramidal events with either in haloperidol vs. flunitrazepam at 90 minutes.\textsuperscript{48}</td>
</tr>
<tr>
<td></td>
<td>Marked sedation</td>
<td>28</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Very few events at 90 minutes with either haloperidol vs. flunitrazepam, no statistical testing reported (3 vs. 3, p=NR).\textsuperscript{48}</td>
</tr>
<tr>
<td></td>
<td>Drug-related adverse events</td>
<td>91</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Very small numbers of events with no statistical testing for detecting differences in droperidol vs. midazolam vs. a combination of droperidol plus midazolam (2 vs. 8 vs. 2, p=NR).\textsuperscript{43}</td>
</tr>
<tr>
<td></td>
<td>Abnormal QT interval</td>
<td>91</td>
<td>Insufficient</td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Very small numbers of abnormal QT intervals with no statistical testing for detecting differences in droperidol vs. midazolam vs. a combination of droperidol plus midazolam (2 vs. 2 vs. 4, p=NR).\textsuperscript{43}</td>
</tr>
<tr>
<td></td>
<td>Medication side effects</td>
<td>0</td>
<td>Insufficient</td>
<td>Low risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No medication side effects reported with either haloperidol plus lorazepam vs. lorazepam.\textsuperscript{64}</td>
</tr>
</tbody>
</table>
Table E. Summary of findings with strength of evidence: Comparative harms of medication protocols for addressing aggressive behavior (Key Question 2b) (continued)

<table>
<thead>
<tr>
<th>Intervention and Comparison</th>
<th>Primary Outcome of Interest</th>
<th>Outcome of Interest</th>
<th>N of Patients Analyzed</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication protocols vs. other medication protocols (continued)</td>
<td>Adverse effects from medication (continued)</td>
<td>Reduction in vital signs</td>
<td>202</td>
<td>Insufficient</td>
<td>High risk of bias, consistency unknown—single study, direct, precise</td>
<td>No significant difference for any reduced vital signs between droperidol vs. lorazepam.68</td>
</tr>
<tr>
<td></td>
<td>Overall treatment-emergent adverse events</td>
<td>Insufficient</td>
<td>558</td>
<td></td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Few overall treatment-emergent adverse events with olanzapine vs. risperidone vs. haloperidol, p=NR.72</td>
</tr>
<tr>
<td></td>
<td>Adverse events considered related to primary antipsychotic medication</td>
<td>Insufficient</td>
<td>558</td>
<td></td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Very few events considered related to primary antipsychotic medication with olanzapine vs. risperidone vs. haloperidol, p=NR.72</td>
</tr>
<tr>
<td></td>
<td>Extra-pyramidal symptoms</td>
<td>Insufficient</td>
<td>558</td>
<td></td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No significant differences (very few events) with olanzapine vs. risperidone vs. haloperidol vs other comparator groups.72</td>
</tr>
<tr>
<td></td>
<td>Discontinuation due to clinically significant adverse events</td>
<td>Insufficient</td>
<td>558</td>
<td></td>
<td>High risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>No significant difference in discontinuation due to clinically significant adverse events with olanzapine vs. risperidone vs. haloperidol vs other comparator groups.72</td>
</tr>
<tr>
<td></td>
<td>Extra-pyramidal events</td>
<td>Insufficient</td>
<td>101</td>
<td></td>
<td>Medium risk of bias, consistency unknown—single study, direct, imprecise</td>
<td>Very small numbers of extrapyramidal events in risperidone vs. olanzapine vs. quetiapine vs. haloperidol, p=0.012.65</td>
</tr>
</tbody>
</table>

N = number; NR = not reported; QT = QRS complex to T wave interval; vs. = versus.
**Key Question 2c: Harms of Strategies To Reduce Seclusion and Restraint Use**

No studies provided information on the comparative harms of staff training, risk assessment, or multimodal, environmental, or medication protocols to reduce seclusion and restraint for patients with active aggression. Thus, all SOE grades were insufficient.

**Key Question 3. Characteristics Modifying the Comparative Benefits or Harms of Strategies**

No studies provided information on how particular characteristics might modify the effectiveness of any of the interventions. Again, SOE is insufficient in all cases.

**Discussion**

Our review aimed to fill gaps in available literature about the comparative effectiveness of various strategies to accomplish one or more of the following goals: prevent aggressive behavior, de-escalate aggressive behaviors, or decrease reliance on seclusion or restraints. An overarching objective, of course, is to improve health outcomes for patients at risk of or exhibiting acute aggressive behavior. We focused on studies in acute care settings.

**Key Findings and Strength of Evidence**

Overall, the evidence base was extremely limited. Of 29 included studies, 18 were pre/post studies. Their inherent high risk of bias precludes drawing inferences of causality, so we did not grade SOE. The main report provides more information on these 18 studies.

We identified 11 studies (mainly RCTs or CRTs) for which we could grade the SOE of one or more outcomes. Of these, 3 were CRTs (for KQ 1); we rated each as medium risk of bias, most commonly because of failure to control either for potential confounding or for intraclass correlations in the CRTs that were eligible for inclusion. No KQs had comparative data supporting an SOE grade that exceeded low strength of evidence. By definition, all findings were of unknown inconsistency (because they are single studies), but all provided direct evidence. In most cases, however, the data reported were imprecise. Thus, we graded these findings as insufficient SOE. In a very small number of cases when data were precise, we graded SOE as low.

Most evidence addressed preventive, unit-wide programs rather than interventions specifically targeting actively aggressive patients; this focus essentially represents the core difference between the CRTs and the RCTs.

Moreover, these analyses could involve samples of patients who were not actively aggressive as well as those who were. These factors prevented us from attributing reduction of aggressive behavior in actively aggressive patients to any particular intervention.

Furthermore, the inexact description of many interventions made it difficult to attribute a change to a particular component. For example, multimodal interventions had components of risk assessment and staff training, and distinguishing between them was sometimes challenging.

As noted earlier, some SOE grades for KQ 1 were low (when we could assign a grade other than insufficient). Findings from eligible studies for KQ 2 were all insufficient, and we had no studies for KQ 3. The variety of measures used to assess aggressive behavior and seclusion and restraint use prevented quantitative synthesis of the meager data that were available.

The table below (Table F) addresses the two studies providing evidence supporting a low SOE, each involving the use of risk assessment protocols to prevent aggressive behavior. Both studies identified lower aggression incidents and rates with use of risk assessment protocols when compared with the usual care conditions. The protocols used had some overlap but differed in important ways. While both trials used the Brøset Violence Checklist as part of the protocol, the van de Sande et al. trial used a more comprehensive protocol that included a Crisis Monitor form and the Kennedy-Axis V (short version) on a daily basis and the full version of the Kennedy-Axis V, the Brief Psychiatric Rating Scale, the Dangerousness Scale, and the Social Dysfunction and Aggression Scale on a weekly basis. The trials also differed in the length of time over which they evaluated their risk assessment protocols. For example, the Abderhalden trial implemented the risk assessment protocol for the first 3 days, whereas the van de Sande et al. trial from The Netherlands used the risk assessment protocol throughout each patient’s hospital stay.

Neither trial analyzed its data in a way that correctly made use of the CRT study design, leading to a risk of bias assessment as medium and, consequently, a low (rather than moderate) SOE rating for the benefit of a risk assessment. We identified no eligible studies assessing the harms of such an intervention.
### Table F. Summary of findings with strength of evidence grades: Comparative benefits of two strategies for preventing aggressive behavior[^a] (*KQ 1a*)

<table>
<thead>
<tr>
<th>Intervention and Comparison/Study Design</th>
<th>Primary Outcome of Interest</th>
<th>Outcome</th>
<th>Strength of Evidence</th>
<th>Supporting Judgment</th>
<th>Findings and Direction of Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk assessment vs. usual care/CRT</td>
<td>Change in aggressive behavior</td>
<td>Aggressive incidents 170 during baseline period, 458 during intervention period</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant 68% RR reduction with risk assessment vs. usual care, p&lt;0.0001 reported; failure to control for intraclass correlations weakens the finding.[^53]</td>
</tr>
<tr>
<td>Rate of severe aggressive incidents 973 post-intervention</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significantly lower risk with structured risk assessment: (RR, 0.59; 95% CI, 0.41 to 0.83); p&lt;0.001 reported; failure to control for intraclass correlations weakens the finding. Decrease achieved since baseline with risk assessment (-41%) vs. usual care (-15%), no statistical testing reported.[^44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in seclusion or restraint</td>
<td>Change in physical attacks 973 post-intervention</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significantly greater decrease with risk assessment (-41%) vs. usual care (7%), p&lt;0.001 reported, failure to control for intraclass correlations weakens the finding.[^44]</td>
<td></td>
</tr>
<tr>
<td>Hours in seclusion 170 during baseline period, 458 during intervention period</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, direct, precise</td>
<td>Significant 45% RR reduction with risk assessment vs. usual care, p&lt;0.0001 reported; failure to control for intraclass correlations weakens the finding.[^53]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in coercive[^b] incidents 973 post-intervention</td>
<td>Low</td>
<td>Medium risk of bias, consistency unknown—single study, indirect, precise</td>
<td>Significant decrease from baseline with risk assessment (-27%) vs. compared with usual care (+10%), p&lt;0.001; failure to control for intraclass correlations weakens the finding.[^44]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[^a]: For KQ 1a, we had no studies of eligible study design for environmental or group psychotherapeutic interventions or multimodal interventions; thus, we could not rate risk of bias.

[^b]: Coercive measures covered a wide range of measures from forced injection of psychotropic medication to seclusion and mechanical restraint.[^44]

[^CI]: confidence interval; CRT= cluster randomized trial; KQ = Key Question; N = number; NR = not reported; RR = relative risk; vs. = versus.
The low confidence in these SOE grades (very few low grades; mainly insufficient grades that are not included in these tables because we had no relevant studies) reflect a critical limitation of the reviewed research. The grades call into question both the reproducibility or replicability and the generalizability of results. Subsequent studies, assuming that they are well designed and take statistical issues accurately into account, are likely to affect these findings substantially, although in what direction remains unclear. Future research, with the same assumptions, may confirm some findings but provide more information that might lead to higher SOE grades.

**Findings in Relationship to What Is Already Known**

This limited body of evidence is consistent with prior findings. Earlier reviews emphasized the lack of high-quality and effective intervention studies to prevent the development of aggressive behavior in acute care settings.\textsuperscript{10,13,73,74} The absence of relevant literature has been similarly reported for patients with actively aggressive behavior, whether alternative strategies were being compared with seclusion and restraints\textsuperscript{10,13} or whether alternatives to seclusion and restraints were being compared with each other.\textsuperscript{10,11,13,73} The lack of literature relevant to comparative harms of these interventions has also been identified.\textsuperscript{75} Our review updates and confirms these findings, although we do include potentially relevant pharmacologic interventions that had not been reported before.

What our review adds is the finding that a general application to all individuals on inpatient psychiatric units (i.e., not just to those who are actively aggressive) of a strategy that involves a risk assessment component may decrease subsequent aggressive behavior. Earlier reviews of risk assessments assessed whether they could decrease agitation, which is often considered a lower-level precursor to aggression. However, both the CRTs that evaluated the effectiveness of risk assessment had data analytic limitations related to using a cluster randomized design. Specifically, investigators had not analyzed their data so as to account appropriately for the clustered nature of the data; this drawback likely affected each trial’s results (e.g., increased the risk of a type I error). Finally, our results can be considered in the context of prior research about the impact of risk assessment practices on patients’ agitation.\textsuperscript{25} Specifically, we identified a potential relationship between using risk assessment and lower aggression in acute care settings (albeit with the statistical limitations we noted); earlier research had found that using risk assessment is associated with reduced agitation.

**Applicability**

The scope of our review encompassed adults with a diagnosed psychiatric disorder, including delirium, in an acute care hospital setting and adults with severe psychiatric symptomatology in an emergency department setting. In addition, we included studies of patients for whom attempts were made to prevent aggressive behavior or to de-escalate that behavior if they became actively aggressive. This focus on acute care settings (rather than psychiatric hospitals, which can involve both acute and longer-term lengths of stay) prevented inclusion of the few otherwise eligible studies that addressed the use of the Six Core Strategies,\textsuperscript{74} a key strategy in widespread use in psychiatric units worldwide. For example, some evidence of psychiatric hospitals with longer stays (i.e., months to years) suggests benefit of multimodal interventions.\textsuperscript{76} We did not include this information because the setting was not an acute one. Indeed, we were surprised that no eligible trials tested application of the Six Core Strategies for decreasing aggressive behavior, given its influence on practices both in the United States and internationally.\textsuperscript{77,79}

The populations and settings in the included studies were relevant to those we were targeting. Mean ages generally ranged from 38 to 40 years. Studies varied widely in the percentages of patients who were male or female. We found little information on other sociodemographic characteristics of patients.

Interventions were in line with clinical practice in acute care units. However, the specifics of how investigators implemented their interventions were not always clear; hence, how to reproduce or replicate them is also uncertain. This point is especially relevant to the multimodal protocols, where varying fidelity to multiple components made it difficult to attribute benefits to specific components.

Studies generally compared interventions with usual care. Usual-care practices appeared to be consistent with standard practice on psychiatric and medical units. The only studies directly comparing alternative strategies with each other involved medication protocols. Only one study compared an alternative strategy (first choice involuntary medication) directly with seclusion (considered usual care in that country).

Outcomes measured were quite diverse; this fact precluded any kind of quantitative synthesis of data. For example, changes in any aggressive incidents versus changes in severe aggressive incidents were not regarded as combinable outcomes. Also, most studies reported short-term, but not long-term, outcomes. One study
reported long-term outcomes such as quality of life, patient experience, and subsequent aggressive behavior. Two studies reported on use of services and economic outcomes.

Nineteen studies addressed individuals on an acute care psychiatric unit (rather than a medical or emergency department setting). Approximately half of the studies were conducted in the United States. However, of the 11 eligible studies, only 4 were from U.S. settings (1 high risk of bias CRT in inpatient psychiatric settings, 69 1 high risk of bias retrospective cohort study addressing delirium in an intensive care unit, 51 and 1 high risk of bias RCT 68 and 1 low risk of bias RCT 64 both addressing aggression in an emergency department). Indeed, 5 of the eligible studies involving inpatient psychiatric settings were conducted in countries other than the United States. The 2 studies forming the basis for the single low SOE intervention, risk assessment, 44,53 were both conducted outside the United States. How substantially clinical practice in sites outside the United States differs from current U.S. practice is not clear. This finding implies that the applicability of findings from outside the United States may be questioned.

Implications for Clinical and Policy Decisionmaking

The paucity of evidence means that most of our implications are for future research rather than clinical or policy judgments. The handful of findings that we graded as low SOE may provide some implications for clinical practice or policy judgments.

In particular, a limited number of risk assessment interventions subsequently led to less aggressive behavior (low SOE) and reduced the subsequent use of seclusion and restraints (low SOE). These findings suggest the need for clinicians to consider carefully the role of these strategies as interventions on psychiatric inpatient units. Specifically, acute care practitioners and administrative staff will need to balance the low SOE with the reality that violence is a pressing (indeed growing) concern and poses significant disruptions to quality of care in such settings. The questions that may arise, for example, include: Is the limited evidence currently available sufficient for evaluating effectiveness? Should implementation decisions be delayed until more evidence becomes available? What is the role of quality measures, designed to create incentives to improve the quality of care, when the evidence base for those measures is unclear?

As to the last question, we are unaware of any ongoing trials that will add to the current sparse body of evidence regarding the benefits of risk assessment protocols. Furthermore, we cannot comment on potential harms or costs associated with implementing risk assessment protocols. Indeed, with no eligible data from U.S. inpatient psychiatric settings, determining how these interventions might be applied in this country and what modifications might be necessary are key next steps.

Research Recommendations

Major evidence gaps exist in this increasingly worrisome clinical arena; they point to important next steps for research in preventing and de-escalating aggressive behavior in acute care settings. The SOE grades informing decisionmaking in this area are minimal. A major void is well-designed, adequately powered, properly analyzed comparative trials that address questions of prevention and de-escalation. The validity of findings from the three reasonably well-designed CRTs was severely limited by analyses that did not properly control for the clustered nature of the data. We applaud the efforts to conduct comparative trials, but this evidence base does not convincingly show the efficacy of most of these strategies; that fact complicates the design of strong comparative studies and reflects a gap that may need to be addressed first.

Head-to-head trials that move beyond a usual-care comparator to examine various interventions against each other are needed to guide decisionmaking. The critical element is identifying the “right” interventions to compare, to make the most efficient use of research time and funding on this topic. More evidence that can speak to differential effectiveness of various interventions would allow clinicians and administrators to balance effectiveness with implementation and resource costs.

Investigators leading trials in the future must clearly describe their interventions. Only in this way can other research teams sensibly try to reproduce or replicate such studies and help confirm which components of the interventions may be the most (or least) effective. Risk assessment strategies, which have some evidence for preventing aggressive behavior, need to be described in more detail to enable them to be compared with each other and allow variations within these approaches to be compared.

Currently, clinicians and investigators do not know the accuracy of risk assessment tools. These are necessary to identify patients at high risk of aggressive behavior and, hence, to develop an effective plan to manage potential or real aggressive behavior. For that reason, more work on documenting the measurement properties of these tools is needed.
All future trials must report on consistently defined and clinically meaningful outcomes, both short term and long term. Selection of these outcomes needs to be informed by key stakeholders, including patients. Crucial short-term outcomes include reliable and valid measures of aggressive behavior and of seclusion and restraint actions. Using well-established, reliable, and valid assessments of aggression that can be harmonized across studies (and, ideally, countries) is crucial, as well, for future systematic reviews on these topics. In addition, research teams should increase adherence to the Consolidated Standards of Reporting Trials (CONSORT) statement regarding the reporting of clinical trials (including CRTs).

Key long-term outcomes must involve more patient-centered outcomes, including, for instance, quality of life or other patient-reported outcomes. Patient perspectives of harms, including treatment preferences, are largely missing from the literature in acute care settings, and this gap should be remedied. Measures of the use of health services are important, as are cost implications and data. Investigators should incorporate implementation factors, such as acceptability, feasibility, and sustainability, into their designs for intervention research in acute care settings.

Available acute care data are almost entirely from inpatient psychiatric settings and from settings outside the United States. In the latter case, standard practices, patient populations, insurance coverage, costs, and various other variables may differ, perhaps considerably. Future well-designed studies of inpatient psychiatric settings need to be conducted in U.S. settings. In addition, informative data must be collected from acute care medical and surgical units and from emergency department settings.

Finally, we had no useful data on modifiers of treatment effectiveness. Thus, future studies (including comparative trials) need to assess how variables such as age and other sociodemographic or economic factors, specific diagnosis (and perhaps coexisting conditions), and specific treatment components modify or mediate the effects of the interventions studied. Consideration of effect modifiers must be powered appropriately, although we acknowledge that in this clinical area, achieving adequate sample sizes for comparative trials of these types of interventions (perhaps apart from medication protocols) may prove challenging.

Conclusions

Given the ethical imperative for treating all patients with dignity, the clinical mandate of finding evidence-based solutions to these mental health challenges, and the legal liability associated with failure to assess and manage violence risk across the treatment continuum, the need for evidence to guide clinical and policy decisionmaking for de-escalating aggressive behavior is critical. This point is particularly true of acute care settings for at least two reasons: comprehensive clinical and violence risk information may not always be readily available in such institutions, and patient management must be balanced against staffing and treatment limitations unique to each individual setting.

The current evidence base leaves clinicians, administrators, policymakers, and patients without clear guidance on how to best prevent and de-escalate aggressive behaviors in acute care settings. Only risk assessment had any reasonable evidence that they can decrease aggression and reduce seclusion and restraint; however, the strength of that evidence was, at best, low. Evidence for de-escalating aggressive behavior is even more limited. More research is needed to guide clinicians, administrators, and policymakers on how to best prevent and de-escalate aggressive behavior in acute care settings.

References


**Full Report**