TECHNICAL COMMENTARY

Benzodiazapines

NeuRA Discover. Conquer. Cure.

SCHIZOPHRENIA LIBRARY

Introduction

Benzodiazepines may be implemented as an adjunct to antipsychotic medication in order to treat acute symptoms of psychosis such as agitation, aggression, irritability, or anxiety. They may also be used to treat side effects of antipsychotic medications such as movement disorders including tardive dyskinesia, however they are associated with their own side effects and are associated with well-established patterns of tolerance and dependence, so they are prescribed with caution.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2000 that report results separately for people with a diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder or first episode schizophrenia. Review were identified by searching the databases MEDLINE, EMBASE, CINAHL, Current Contents, PsycINFO and the Cochrane library. Hand searching reference lists of identified reviews was also conducted. When multiple copies of reviews were found, only the most recent version was included. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[1]. Reviews reporting less than 50% of items have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large or if there is a dose dependent response. We have also taken into account sample size and whether results are consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found four systematic reviews that met our inclusion criteria[3-6].

- Moderate to low quality evidence finds no benefit of adjunctive benzodiazepines for agitation or excitation in the short term (up to 30 minutes); however benzodiazepines may be associated with a faster time to sedation than antipsychotics alone, better global improvements (up to 60 minutes) and lower risk of movement disorders. No benefit was found between adjunct benzodiazepines and antipsychotics alone for other outcomes.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Bergman H, Bhoopathi PS, Soares-Weiser K

Benzodiazepines for antipsychotic-induced tardive dyskinesia

Cochrane Database of Systematic Reviews 2018; 1: CD000205

View review abstract online

Comparison	**Diazepam vs. placebo**
Summary of evidence	**Moderate to low quality evidence (very small samples, imprecise, consistent, direct) suggests no differences between groups.**
Tardive dyskinesia	
No significant differences between groups; Clinically meaningful improvement: 2 RCTs, N = 32, RR = 1.12, 95%CI 0.60 to 2.09, p = 0.72, I^2 = 14%, p = 0.28 Deterioration: 2 RCTs, N = 30, RR = 1.48, 95%CI 0.22 to 9.82, p = 0.69, I^2 = 19%, p = 0.27	
Risks	No differences in adverse effects.
Consistency in results‡	Consistent
Precision in results§	Imprecise
Directness of results‖	Direct

Dold M, Li C, Tardy M, Khorsand V,Gillies D, Leucht S

Benzodiazepines for schizophrenia

CochraneDatabase of Systematic Reviews 2012; Issue 11. Art. No.: CD006391. DOI: 10.1002/14651858.CD006391.pub2

View review abstract online

Comparison	**Benzodiazepines and antipsychotics (varying) vs. placebo and antipsychotics (varying). Treatment duration 24 hours to 10 weeks.**
Summary of evidence	**Moderate to low quality evidence (very small samples, precise,**

	inconsistent, direct) suggests benzodiazepines may improve global state for up to 60 minutes when compared to placebo, but not over the longer term. Movement disorder may be lower and somnolence higher with benzodiazepines. There was no benefits for study retention, mental state, aggression, and hospitalisation.
Global state	
1 small RCT reported a significant, small to medium-sized effect for improved global state at 30 minutes in people receiving adjunct benzodiazepines, with a large, non-significant trend effect by 60 minutes and no differences at 12 hours or 3 weeks; 30 minutes: 1 RCT, N = 45, RR = 0.38, 95%CI 0.18 to 0.80, *p* = 0.011 60 minutes: 1 RCT, N = 45, RR = 0.07, 95%CI 0.00 to 1.13, *p* = 0.061 12 hours: 1 RCT, N = 67, RR = 0.96, 95%CI 0.68 to 1.34, *p* = 0.80 ≥3 weeks: 6 RCTs, N = 511, RR = 0.87, 95%CI 0.49 to 1.54, *p* = 0.62, I^2 = 68%, *p* = 0.01	
Mental state	
No significant differences between groups; Modified BPRS (psychosis specific items) After 1 hour: 1 RCT, N = 67, WMD = -4.00, 95%CI -9.51 to 1.51, *p* = 0.15 After 12 hours: WMD = 2.00, 95%CI -3.27 to 7.27, *p* = 0.46. BPRS After 3 days: 1 RCT, N = 28, WMD = 6.81, 95%CI 0.32 to 13.30, *p* = 0.04 (favoured control group) After 28 days:1 RCT, N = 61, WMD = -4.50, 95%CI -14.93 to 5.93, *p* = 0.40 After 8 weeks:1 RCT, N = 97, WMD = -3.20, 95%CI -6.54 to 0.14, *p* = 0.061 PANSS After 2 weeks: 1 RCT, N = 80, WMD = -6.20, 95%CI -12.55 to 0.15, *p* = 0.056 (trend favoured adjunctive treatment)	
Sedation and aggressive behaviour	
Significant, medium-sized effects suggest greater sedation in the benzodiazepine group at 30 minutes, reducing to a small effect by 60 minutes, with no difference in rate of sleeping at 12 hours Tranquilised at 30 minutes: 1 RCT, N = 45, RR = 2.25, 95%CI 1.18 to 4.30, *p* = 0.014 Tranquilised at 60 minutes: 1 RCT, N = 45, RR = 1.39, 95%CI 1.06 to 1.83, *p* = 0.019 Asleep at 12 hours: 1 RCT, N = 67, RR = 0.85, 95%CI 0.51 to 1.41, *p* = 0.53 *No significant difference in aggressive behaviour;*	

At 1 hour: 1 RCT, N = 67, WMD = -3.00, 95%CI -8.27 to 2.27, *p* = 0.26 At 12 hours: 1 RCT, N = 67, WMD = 0.00, 95%CI -5.27 to 5.27, *p* = 1.0	
Service use	
No significant differences between groups in the likelihood of hospital discharge by 3 days; 1 RCT, N = 28, RR = 1.25, 95%CI 0.42 to 3.70, *p* = 0.69	
Leaving the study early	
No significant differences between groups; Up to 10 weeks: 16 RCTs, N = 1185, RR = 1.36, 95%CI 0.81 to 2.30, *p* = 0.25, I^2 = 0% Up to 10 weeks due to side effects: 6 RCTs, N = 413, RR = 3.24, 95%CI 0.68 to 15.45, *p* = 0.14, I^2 = 0% Up to 10 weeks due treatment inefficacy: 6 RCTs, N = 347, RR = 0.76, 95%CI 0.17 to 3.42, *p* = 0.72, I^2 = 0%	
Risks	Use of antiparkinson medication was significantly lower in the benzodiazepine group (N = 282, RR 0.68, 95%CI 0.49 to 0.94, *p* = 0.021, I^2 = 12%), and there was a trend for lower dystonia (N = 155, RR 0.32, 95%CI 0.08 to 1.22, *p* = 0.09, I^2 = 0%) and parkinsonism (N = 211, RR 0.26, 95%CI 0.07 to 1.06, *p* = 0.06, I^2 = 0%). Somnolence was significantly higher in the benzodiazepine group (N = 118, RR 3.30, 95%CI 1.04 to 10.40, *p* = 0.042, I^2 = 0%). There was no significant differences (*p* > 0.05) in risk of: cardiovascular reaction (N = 144, RR 1.53, 95%CI 0.19 to 12.11, I^2 = 0%), ataxia (N = 127, RR 1.26, 95%CI 0.27 to 5.87, I^2 = 20%), anorexia (N = 60, RR 0.33, 95%CI 0.01 to 7.87), allergic reaction (N = 84, RR 1.36, 95%CI 0.24 to 7.75), blurred vision (N = 144, RR 0.96, 95%CI 0.10 to 9.04, I^2 = 0%), confusion (N = 60, RR 0.33, 95%CI 0.01 to 7.87), depression (N = 60, RR 1.01, 95%CI 0.07 to 15.26), diarrhoea (N = 60, RR 3.00, 95%CI 0.13 to 70.83), dizziness (N = 257, RR 1.96, 95%CI 0.88 to 4.37, I^2 = 2%), drowsiness (N = 151, RR 0.52, 95%CI 0.13 to 2.05, I^2 = 0%), excitation (N = 60, RR 1.45, 95%CI 0.26 to 8.11), gastrointestinal reaction (N = 84, RR 0.30, 95%CI 0.01 to 7.25), headache (N = 142, RR 0.78, 95%CI 0.04 to 14.93, I^2 = 45%), dry mouth (N = 269, RR 1.63, 95%CI 0.38 to 7.87, I^2 = 0%), increased salivation (N = 144, RR 1.98, 95%CI 0.26 to 14.94, I^2 = 0%), insomnia (N = 144, RR 1.50, 95%CI 0.41 to 5.44, I^2 = 0%), lactation (N = 84, RR 0.30, 95%CI 0.01 to 7.25), restlessness (N = 118, RR 1.16[0.40 to 3.36], I^2 = 0%), sensory disturbance (N = 84, RR 0.91, 95%CI 0.06 to 14.06), sleep disorder (N = 58, RR 0.41, 95%CI 0.02 to 9.60), vomiting (N = 60, RR 3.00, 95%CI 0.13 to

	70.83).
Consistency in results	Consistent for all except clinical improvement at 3 weeks. Not applicable for outcomes with 1 study.
Precision in results	Imprecise, unable to assess WMD.
Directness of results	Direct

Gillies D, Beck A, McCloud A, Rathbone J

Benzodiazepines for psychosis-induced aggression or agitation

Cochrane Database of Systematic Reviews 2005; Issue 4. Art. No.: CD003079. DOI: 10.1002/14651858.CD003079.pub2.

View review abstract online

Comparison 1	**Benzodiazepines plus antipsychotics (lorazepam + haloperidol), of 1-6 doses over up to 8 hours vs. benzodiazepines alone (lorazepam) for people with acute psychosis. Primarily administered intramuscularly.**
Summary of evidence	**Moderate to low quality evidence (very small samples, consistent, imprecise, direct) suggests no differences between groups in the need for additional medication or for extrapyramidal side effects. Low quality evidence from individual studies with very small samples are unable to determine any benefits for study retention, global state, behaviour or other adverse effects.**

Global state

No significant differences between groups;

Need for additional medication (up to 48 hours after admission): 2 RCTs, N = 83, RR = 1.02, 95%CI 0.79 to 1.32, $p = 0.90$, Q = 1.25, $p = 0.26$, $I^2 = 20\%$

Clinical global improvement (CGI, up to 1 hour): 1 RCT, N = 20, RR = 1.47, 95%CI 0.66 to 3.25, $p = 0.35$

1 RCT (N = 20) reported that no participants in either group left the study early (up to 48 hours)

Behaviour

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

No significant differences between groups; Behavioural improvements in using Overt Aggression Scale (OAS) (up to 1 hour): 1 RCT, N = 20, RR = 0.11, 95%CI 0.01 to 1.74, *p* = 0.12	
Risks	2 RCTs found no significant differences in the risk of extrapyramidal effects, (N = 83, RR = 1.94, 95%CI 0.18 to 20.30, *p* = 0.58, I^2 = 0%). One RCT found no significant difference (*p* > 0.05) between groups (N = 63) for risk of ataxia (RR 1.45, 95%CI 0.26 to 8.11), dizziness (RR 0.65, 95%CI 0.12 to 3.61), dry mouth (RR 0.58, 95%CI 0.15 to 2.23), or speech disorder (RR 1.45, 95%CI 0.26 to 8.11).
Consistency in results	Consistent where applicable (need for additional medication and extrapyramidal side effects).
Precision in results	Imprecise
Directness of results	Direct
Comparison 2	**Benzodiazepines plus antipsychotics (lorazepam or alprazolam + haloperidol) vs. antipsychotics alone (haloperidol) for people with acute psychosis. Both oral and intramuscular medications reported.**
Summary of evidence	**Moderate to low quality evidence (very small samples, imprecise, consistent, direct) suggests those in the benzodiazepine plus haloperidol group showed significantly lower risk of extrapyramidal side effects. No differences between groups were reported for need for additional medication, mental state, hospitalisation, or other adverse effects.**
Need for additional medication	
No significant differences between groups; Need for additional medication (up to 48 hours): 1 RCT, N = 67, RR = 0.95, 95%CI 0.79 to 1.15, *p* = 0.62	
Mental state	
No significant differences between groups; BPRS-total: 1 RCT, N = 28, WMD = 0.01, 95%CI -7.26 to 7.28, *p* = 1.0 BPRS-psychosis subscale: 1 RCT, N = 28, WMD = -1.93, 95%CI -5.73 to 1.87, *p* = 0.32	
Hospitalisation – likelihood of early discharge	

No significant differences between groups; 1 RCT, N = 28, RR = 0.90, 95%CI 0.54 to 1.50, *p* = 0.69	
Risks	From 2 RCTs, medium effect size suggests the risk of extrapyramidal effects was significantly lower in the benzodiazepine + antipsychotic group (N = 95, RR = 0.45, 95%CI 0.22 to 0.94, *p* = 0.035, I^2 = 0%), but there was no difference in need for anticholinergic medication reported in 1 RCT (N = 28, RR = 0.56, 95%CI 0.25 to 1.24, *p* = 0.15). One RCT (N = 67) reported no significant differences (*p* > 0.05) between groups in rates of ataxia (RR 3.28, 95%CI 0.36 to 29.97), dizziness (RR 0.73, 95%CI 0.13 to 4.09), dry mouth (RR 1.09, 95%CI 0.24 to 5.04), or speech disorder (RR 0.82, 95%CI 0.20 to 3.39).
Consistency in results	Consistent where applicable (extrapyramidal side effects).
Precision in results	Precise for additional medication, imprecise for adverse effects. Unable to assess WMD as standardised measure not reported.
Directness of results	Direct

Zeller SL, Rhoades RW

Systematic reviews of assessment measures and pharmacologic treatments for agitation

Clinical Therapeutics 2010; 32(3): 403-425

View review abstract online

Comparison	**Benzodiazepines (lorazepam) plus antipsychotics (risperidone, haloperidol, olanzapine, aripiprazole) for agitation in acute psychosis. Both oral and intramuscular medications reported.**
Summary of evidence	**Moderate to low quality evidence (mixed sample sizes, unable to assess consistency or precision, direct) suggests no benefit of adjunctive lorazepam over antipsychotic treatment for mental state (PANSS scores) in the immediate to short-term.**
Agitation/Excitation	
One trial (N = 162) compared lorazepam as an adjunct to either haloperidol or risperidone, over 24 hours. Both groups showed significant improvement in PANSS-excitation scores (*p* < 0.001), with	

	no difference between groups. Two trials (N = 704) in people with schizophrenia spectrum diagnoses compared lorazepam as an adjunct to haloperidol, olanzapine, or aripiprazole, for between 5 days and 3 weeks. Both groups showed significant improvement in PANSS scores on excitation or agitation subscales ($p < 0.001$), with no difference between groups. One small trial (N = 20) compared adjunct lorazepam with either risperidone or haloperidol over 90 minutes, and reported no differences between groups in PANSS scores. One trial (N = 148) compared adjunct lorazepam with different doses of olanzapine over 24 hours. Both groups showed significant improvement in PANSS-excitation scores ($p < 0.001$), with no difference between groups. One trial (N = 20) compared adjunct lorazepam and haloperidol with lorazepam alone over 180 minutes. The haloperidol group was associated with significantly lower aggression at 60 minutes (p = 0.04). One trial (N = 98) compared adjunct lorazepam and haloperidol to either lorazepam or haloperidol alone, over 24 hours. Adjunct therapy showed greater improvements in Agitation scale over lorazepam alone, but no difference to haloperidol alone.
Risks	There was no apparent difference between groups for adverse effects (*p* value not reported). Reported side effects included somnolence, headache, agitation and hyperkinesia, pain, insomnia, dizziness, nervousness, and dry mouth.
Consistency in results	No measure of consistency is reported.
Precision in results	No measure of precision is reported.
Directness of results	Direct

Explanation of acronyms

BPRS = Brief Psychiatric Rating Scale, CGI = clinical global improvement scale, CI = Confidence Interval, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), N = number of participants, NA = Not applicable, OAS = Overt Aggression Scale, *p* = statistical probability of obtaining that result ($p < 0.05$ generally regarded as significant), PANSS = Positive and Negative Syndrome Scale, Q = Q statistic for the test of heterogeneity, RCT = Randomised Controlled Trial, RR = Relative Risk, vs = versus, WMD = Weighted Mean Difference

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[7].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardised mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) which allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. Less than 0.4 represents a small effect, around 0.5 a medium effect, and over 0.8 represents a large effect[7].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[8]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[7];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[9].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C, which allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Gillies D, Beck A, McCloud A, Rathbone J (2005): Benzodiazepines for psychosis-induced aggression or agitation. *Cochrane Database of Systematic Reviews*.
4. Zeller SL, Rhoades RW (2010): Systematic reviews of assessment measures and pharmacologic treatments for agitation. *Clinical Therapeutics* 32: 403-25.
5. Dold M, Li C, Tardy M, Khorsand V, Gillies D, Leucht S (2012): Benzodiazepines for schizophrenia. *Cochrane Database of Systematic Reviews* 11: CD006391.
6. Bergman H, Bhoopathi PS, Soares-Weiser K (2018): Benzodiazepines for antipsychotic-induced tardive dyskinesia. *Cochrane Database of Systematic Reviews* 1: CD000205.
7. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
8. Sommer IE, van Westrhenen R, Begemann MJH, de Witte LD, Leucht S, Kahn RS (2014): Efficacy of Anti-inflammatory Agents to Improve Symptoms in Patients With Schizophrenia: An Update. *Schizophrenia Bulletin* 40: 181-91.
9. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*