A meta-analysis of neuropsychological change to clozapine, olanzapine, quetiapine, and risperidone in schizophrenia

Neil D. Woodward1, Scot E. Purdon2, Herbert Y. Meltzer3 and David H. Zald1

1 Department of Psychology, Vanderbilt University, Nashville, TN, USA
2 Department of Psychiatry, University of Alberta, Edmonton, Alberta, Canada
3 Department of Psychiatry, Vanderbilt University Medical Center, Nashville, TN, USA

Abstract

Cognitive impairment is a core feature of schizophrenia and a major impediment to social and vocational rehabilitation. A number of studies have claimed cognitive benefits from treatment with various atypical antipsychotic drugs (APDs). The currently available evidence supporting cognitive improvement with atypical APDs was evaluated in two meta-analyses. Studies that (1) prospectively examined cognitive change to the atypical APDs clozapine, olanzapine, quetiapine, and risperidone, (2) included a commonly used neuropsychological test, and (3) provided data from which relevant effect sizes could be calculated, were included. Forty-one studies met these criteria. Neuropsychological test data from each study were combined into a Global Cognitive Index and nine cognitive domain scores. Two meta-analyses were carried out. The first included 14 controlled, random assignment trials that assigned subjects to an atypical APD and a typical APD control arm. The second analysis included all prospective investigations of atypical treatment and the within-group change score divided by its standard deviation served as an estimate of effect size (ES). The first analysis revealed that atypicals are superior to typicals at improving overall cognitive function (ES = 0.24). Specific improvements were observed in the learning and processing speed domains. The second analysis extended the improvements to a broader range of cognitive domains (ES range = 0.17–0.46) and identified significant differences between treatments in attention and verbal fluency. Moderator variables such as study blind and random assignment influence results of cognitive change to atypical APDs. Atypical antipsychotics produce a mild remediation of cognitive deficits in schizophrenia, and specific atypicals have differential effects within certain cognitive domains.

Received 14 July 2004; Reviewed 28 September 2004; Revised 21 October 2004; Accepted 27 October 2004

Key words: Atypical antipsychotics, meta-analysis, neuropsychology, schizophrenia.

Introduction

Cognitive dysfunction is fundamental to schizophrenia (Bleuler, 1950; Kraepelin and Robertson, 1919) and readily demonstrated on a variety of neuropsychological instruments (Kolb and Whishaw, 1983). Patients with schizophrenia typically perform one to two standard deviations below normal on a variety of measures, especially those that assess executive functions, verbal skills, processing speed, and attention (Bilder et al., 2000; Fuller et al., 2002; Heinrichs and Zakzanis, 1998; Hoff et al., 1992; Saykin et al., 1994). Cognitive impairment in schizophrenia relates directly to socio-vocational functioning (Green, 1996; Green et al., 2000), and exerts a greater influence on functional outcome than the presence or severity of the positive or negative symptoms of schizophrenia (Velligan et al., 2000). Furthermore, associations between particular cognitive skills and specific dimensions of outcome have been articulated. Thus, the relationships between cognitive impairments and psychosocial deficits may provide a basis for the prediction of functional changes that should result from treatment-specific changes in cognitive status.

After many years of null results with typical antipsychotic drugs (APDs), and an early negative study of the effect of clozapine on cognition (Goldberg et al., 1993), a series of studies identified significant improvements in cognition with other atypical APDs in addition to clozapine (Bilder et al., 2002; Galletly et al., 1999; Hagger et al., 1993; Meltzer and McGurk, 1999;
As will be discussed, the cognitive enhancement reported in these early studies could have been artifacts related to repeated testing, study characteristics, or other potential biases. Alternatively, the apparent cognitive enhancement may be related to one or more of the following effects of the atypical APDs which are not shared by typical APDs: (1) increased release of dopamine (DA) and acetylcholine (ACh) in the prefrontal cortex and hippocampus (Ichikawa et al., 2002; Kuroki et al., 1999; Parada et al., 1997; Shirazi-Southall et al., 2002); (2) antagonism of 5-HT2A, 5-HT2C or 5-HT6 receptors (Meltzer, 1999); and (3) stimulation of 5-HT1A receptors (Ichikawa et al., 2001). Increased release of DA may lead to stimulation of D1 and D3 receptors, in particular, which might have a beneficial effect on cognition, assuming that these receptors are under-stimulated in schizophrenia. Increased release of ACh might lead to enhancement of M1, M4, or a2 nicotinic acid post-synaptic receptors, all of which have been suggested to be involved in cognitive impairment in schizophrenia (Bymaster et al., 2003; Olincy et al., 1997; Simosky et al., 2003). The atypical APDs also differ from one another in their relative actions on these systems. Clozapine is an M1 and M4 agonist, an effect which other atypical APDs lack (Olianas et al., 1999; Zorn et al., 1994). Blockade of M2 receptors by clozapine or olanzapine in vivo would be expected to increase the release of ACh. Stimulation of M1 and M4 receptors has been shown to improve memory and learning in animal models (Felder et al., 2001). Risperidone has a relatively high affinity and long dissociation latency period for D2 receptors (Kapur and Seeman, 2001; Lavallée et al., 1999; Seeman, 2002), suggesting that patients receiving risperidone may be more likely to display adverse effects associated with DA antagonism in the striatum including greater extrapyramidal symptoms (EPS) and reduced procedural learning, especially with doses above 6 mg/d. A recent meta-analysis of EPS prevalence in clinical trials and preliminary evidence of reduced procedural learning with risperidone, relative to clozapine and olanzapine, provides support for this prediction (Bedard et al., 2000; Leuch et al., 1999; Purdon et al., 2003). Thus, there are not only neurochemical reasons to expect atypical APDs to improve cognitive function, relative to typical APDs, but differences between treatments within the atypical APD class might also be anticipated.

The significant methodological differences that exist across studies undermine attempts to draw definitive conclusions on the efficacy and differential benefits of atypical APDs to cognition in schizophrenia. Two earlier quantitative reviews of published studies up to 2000 identified significant gains with atypical APDs in several cognitive domains including verbal fluency, vigilance and selective attention, secondary memory, and visuomotor skills (Harvey and Keefe, 2001; Keefe et al., 1999). Effect sizes, in terms of Cohen’s d, were typically within the range of 0.20–0.40 suggesting that the improvements may be mild relative to the magnitude of the cognitive deficits seen in patients with schizophrenia. However, the earlier reviews were hampered by the relatively small number of studies that had been carried out prior to 2000, the limited availability of data on olanzapine, and the absence of data on quetiapine. Since the earlier reviews, the results of over 20 studies involving atypical APDs including several large-scale NIMH and industry-sponsored clinical trials have been released and there is now a substantial pool of data on olanzapine’s effects on cognition and results from several investigations of quetiapine (Bilder et al., 2002; Harvey et al., 2003; Purdon et al., 2001b; Velligan et al., 2002).

The large number of studies that have been reported since 2000 make it feasible to examine the effects of relevant methodological characteristics, such as medication blind, random assignment of subjects, and study duration. Earlier reviews have stressed the importance of controlling for these variables to protect against experimenter bias and demand characteristics. However, quantitative comparisons between studies that included these design features and those that did not are lacking. Additional study variables that may be relevant include baseline medication status and medication dosage used in typical control arms. Several investigators have speculated that the cognitive improvements observed with atypical APDs may reflect an avoidance of potentially deleterious effects associated with typical APD treatments rather than a novel enhancement of cognition (Carpenter and Gold, 2002). Definitive support for this contention is lacking although recent investigations suggest that haloperidol may indeed interfere with specific cognitive skills such as processing speed and procedural learning (Bedard et al., 1996, 2000; Blyer and Gold, 2000; Purdon et al., 2002, 2003; Sharma and Harvey, 2000; Stevens et al., 2002). In the case of within-subjects switch studies, the absence of an unmedicated baseline assessment does not rule out the possibility that the improvements observed following a switch to an atypical APD treatment reflect a release from the adverse effects associated with a typical APD rather than a benefit of atypical APD treatment.
The larger number of studies now available for review also permits a more thorough investigation of the unique cognitive benefits for each medication and a preliminary examination of potential differences between them. Although several investigations have directly compared medications within the atypical APD class, with few exceptions (Harvey et al., 2003), interpretation of the results have been limited by the small number of subjects included in treatment groups (Bilder et al., 2002; Purdon et al., 2000). By quantitatively analysing effects across studies, meta-analysis can overcome these sample-size limitations, and help identify possible differences between treatments that may warrant further investigation in clinical trials.

At present, over 40 studies have reported on the effects of clozapine, olanzapine, risperidone andquetiapine on a wide range of neuropsychological tests. The studies were entered into a meta-analysis to: (1) evaluate and extend the findings of the earlier meta-analyses; (2) identify any differences between atypical APD medications on cognitive processes; and (3) identify study characteristics that might be relevant to demonstrations of cognitive change.

Methods

Literature search

Relevant articles were identified through extensive literature searches of computerized databases including Medline, PsycInfo, and Dissertation Abstracts. Key search terms included Schizophrenia, Cognition, Neuropsychology, Neurocognition, Clozapine, Olanzapine, Risperidone, and Quetiapine. In addition, the bibliographies of several earlier reviews were examined (Harvey and Keefe, 2001; Keefe et al., 1999; Meltzer and McGurk, 1999; Purdon, 1999, 2000).

Studies were included in the meta-analysis if they met the following criteria: (1) inclusion of patients with a diagnosis of schizophrenia or schizoaffective disorder as outlined in DSM-III, DSM-III-R, DSM-IV, or ICD-9, ICD-10; (2) prospective study design with a baseline assessment and at least one follow-up assessment; (3) trial duration of at least 1 wk; (4) no antipsychotics, except for the study medications were administered; (5) a baseline sample size of at least 10; (6) results of neuropsychological change to treatment were reported for at least one of the common tests listed in Table 1; and (7) the study was published or ‘in press’ in a peer-reviewed journal as of April 2004. Investigations of geriatric, adolescent (age <18 yr), or high-risk populations were not included. Studies included in the meta-analysis are listed in Table 2.

<table>
<thead>
<tr>
<th>Table 1. Neuropsychological tests and cognitive domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviation</td>
</tr>
<tr>
<td>Vigilance and Selective Attention</td>
</tr>
<tr>
<td>Continuous Performance Test</td>
</tr>
<tr>
<td>Stroop Test</td>
</tr>
<tr>
<td>Trailmaking A</td>
</tr>
<tr>
<td>Working Memory</td>
</tr>
<tr>
<td>Verbal Working Memory*</td>
</tr>
<tr>
<td>Spatial Working Memory*</td>
</tr>
<tr>
<td>Learning</td>
</tr>
<tr>
<td>Paragraph Recall Test/WMS-R/III Logical Memory I</td>
</tr>
<tr>
<td>Verbal List Learning tests</td>
</tr>
<tr>
<td>(learning trials)</td>
</tr>
<tr>
<td>Rey Design Learning Test</td>
</tr>
<tr>
<td>Rey Complex Figure/WMS-R/III Visual Reproduction I</td>
</tr>
<tr>
<td>Cognitive Flexibility and Abstraction</td>
</tr>
<tr>
<td>Wisconsin Card Sorting Test</td>
</tr>
<tr>
<td>WAIS-R/III Similarities</td>
</tr>
<tr>
<td>Processing Speed</td>
</tr>
<tr>
<td>WAIS R/III Digit Symbol</td>
</tr>
<tr>
<td>Substitution/ Digit Symbol</td>
</tr>
<tr>
<td>Modalities Test</td>
</tr>
<tr>
<td>Trailmaking B</td>
</tr>
<tr>
<td>WISC-R Maze Subtest</td>
</tr>
<tr>
<td>Verbal Fluency</td>
</tr>
<tr>
<td>Controlled Oral Word Test</td>
</tr>
<tr>
<td>Category Instance Generation Test</td>
</tr>
<tr>
<td>Visuospatial Processing</td>
</tr>
<tr>
<td>WAIS-R/III Block Design</td>
</tr>
<tr>
<td>Complex Figure Test (copy)</td>
</tr>
<tr>
<td>Tests of Visual Organization*</td>
</tr>
<tr>
<td>Motor Skills</td>
</tr>
<tr>
<td>Finger Tapping Test</td>
</tr>
<tr>
<td>Grooved Pegboard Test/ PIN Test</td>
</tr>
<tr>
<td>Delayed Recall</td>
</tr>
<tr>
<td>Paragraph Recall Test/WMS-R/III Logical Memory II</td>
</tr>
<tr>
<td>Verbal List Learning tests</td>
</tr>
<tr>
<td>(delayed free recall)</td>
</tr>
<tr>
<td>Rey Complex Figure (delayed)/WMS-R/III Visual Reproduction II</td>
</tr>
</tbody>
</table>

* See text for additional information.

Coding of study characteristics

Studies were coded for author and year of publication, corporate sponsorship, schizophrenia subtype classification, baseline medication status, medication blind, random assignment, trial medications, total subjects
Neuropsychological tests and domains

Similar to other meta-analyses of cognition in schizophrenia (Harvey and Keefe, 2001; Heinrichs and Zakzanis, 1998) effect sizes were calculated for individual neuropsychological tests, although in several cases highly similar tests were combined into a single measure (e.g. verbal list learning). These effect sizes were then combined into nine domains, as listed in Table 1, by averaging effect sizes within studies across tests that putatively tap similar skills. A Global Cognitive Index was also created by either averaging all domain effect sizes within a study or using Global Cognitive Index scores in cases where studies reported them. Thus, each study contributed one Global Cognitive Index score and at least one domain effect size. The construction of the domains reported here was based upon prior reviews and earlier studies that utilized large cognitive batteries, contemporary neuropsychological domain constructs, and cognitive domains relevant to outcome in schizophrenia (Bilder et al., 2000, 2002; Green et al., 2000, 2002; Harvey and Keefe, 2001; Heaton et al., 2001; Purdon et al., 2000, 2001b).

The Vigilance and Selective Attention domain included the Continuous Performance/Attention Test, Stroop Test (colour-word score), and Trailmaking A Test.

The Working Memory domain consisted of tests of verbal or spatial working memory. These included the verbal working memory tests Digit Span, Digit Span Distraction, Paced Auditory Serial Addition, Letter-Number Span, and Consonant Trigrams and spatial working memory tests such as the Visual Span subtest of the WAIS-R/III and the Spatial Working Memory Test (Keefe et al., 1995).

The Learning domain included the Rey Serial Design Learning Test (RDLT), paragraph recall tests (WMS-R/III Logical Memory I or the Story Recall Test), verbal list learning tests (California, Crawford, Hopkins or Rey Verbal Learning tests, or the Bushcke Selective Reminding Test), and visual reproduction tests (WMS-R/III Visual Reproduction subtest, the Rey-Osterrieth/Taylor Complex Figure Test (RCFT), or the Benton Visual Retention Test).

The Cognitive Flexibility and Abstraction domain consisted of the Wisconsin Card Sorting Test (perseverative errors or percent perseverative errors score) and the WAIS-R/III Similarities subtest. Timed motor tests occasionally considered to tap executive function (e.g. Trailmaking B) were not included because differential effects of typical and atypical APDs on motor speed might have unduly influenced effect sizes for this domain.

The Processing Speed domain included the WAIS-R/III Digit Symbol Coding or Digit Symbol Modalities
explicitly reported, Hedges’ group means and standard deviations were not divided by their pooled standard deviation. Where cal APD and typical APD groups at study end-point, appropriate alternative methods based on Hedges’ means and standard deviations were used to calculate randomly assigned patients to treatment. Post-treatment comparisons of typical APDs and atypical APDs that ran-

The first analysis included only reports from com-

Calculation of effect sizes and data analysis

Typically, meta-analyses only include controlled studies that randomly assigned subjects to either a control group or an active treatment group. However, this approach would overlook a substantial body of evidence from single-sample studies that may be relevant to the demonstration of cognitive change from atypical APD treatments. In an attempt to preserve scientific rigour without omitting potentially important results, two analyses were undertaken, the first with a conservative approach to the published literature and the second with less conservative restrictions.

Analysis 1

The first analysis included only reports from comparisons of typical APDs and atypical APDs that randomly assigned patients to treatment. Post-treatment means and standard deviations were used to calculate Hedges’ g, the difference between the means of atypical APD and typical APD groups at study end-point, divided by their pooled standard deviation. Where group means and standard deviations were not explicitly reported, Hedges’ g was calculated using appropriate alternative methods based on t or F statistics (Rosenthal, 1994). Where the t or F statistics were also not reported, data were solicited from the original study authors. A weighted average effect-size estimate was calculated for the Global Cognitive Index and each domain by combining data from all studies that examined cognitive change to clozapine, olanzapine, risperidone, or quetiapine. In cases where a study included more than one atypical APD arm, in addition to a typical APD control, or multiple dosing arms, the atypical APD arms were treated as separate samples and effect sizes for each arm were calculated. Effect sizes were combined according to the fixed-effects model (Shadish and Haddock, 1994). Briefly, each effect size was weighted by the inverse of its variance such that effect sizes calculated from studies with larger sample sizes contributed more to the overall effect size when combined. A weighted average effect size, with positive values indicating improvement and negative values indicating a decline in performance, and a corresponding Z statistic was calculated to determine if the weighted average effect size was significantly greater than zero. Given the large number of Z tests carried out, a Bonferroni correction was applied to the critical α. For the domains, the critical α was p = 0.006. In addition, a 95% confidence interval (CI) was calculated for the global and domain effect sizes. To assess the relevance of predefined moderator variables, a measure of effect size homogeneity, the Q statistic, was also calculated for each neuropsychological domain and the Global Cognitive Index (Hedges and Vevea, 1998). The Q statistic has a χ² distribution with k – 1 degrees of freedom, where k is the number of effect sizes being combined. The critical α for the Q statistic was set at 0.05. When the assumption of homogeneity was rejected the effect sizes were combined using the random-effects model. In the moderator variable analysis, the Q statistic was partitioned into a between-groups component, Q_BET, and a within-groups component, Q_W (analogous to a one-way ANOVA). A moderator variable was considered significant if it effectively separated the effect sizes into separate categories (i.e. Q_BET was significant) that did not have significant within-group variation (i.e. Q_W was not significant). The R² value was also calculated for each significant moderator variable to assess the strength of the relationship between moderator and dependent variables. Moderator variables included the coded study characteristics of baseline medication status (typical APDs vs. unmedicated), schizophrenia subtype classification (early phase combined with general, vs. treatment refractory), and corporate sponsorship of study (yes vs. no). In addition, correlations between effect sizes and the continuous variables haloperidol arm dose at study end-point and study duration were carried out. To avoid violations of independence in the moderator variable analysis, average effect sizes were calculated across groups for the three studies that examined cognitive change in more than one atypical treatment or dosing arm (Bilder et al., 2002; Purdon et al., 2000; Velligan et al., 2002) and for four risperidone studies.
that reported results from the same trial (Green et al., 1997; Kern et al., 1998, 1999; McGurk et al., 1997).

Analysis 2

The second analysis included all prospective studies, regardless of whether or not participants were randomly assigned to treatment, including single-sample switch studies. Investigations of cognitive change following a shift from one atypical APD to another were not included. A single sample, within groups, repeated measures index of effect size, the mean change score divided by its standard deviation, analogous to Hedges’ $g$ was used as the estimate of effect size (Rosenthal, 1994). It should be noted that this method for calculating effect sizes probably yields different effect sizes than those reported in Analysis 1 since each group acts as its own control in a repeated-measures design. Thus, comparisons between Analyses 1 and 2 should not be made. Paired $t$ tests or alternative repeated-measures values were available to calculate an effect size for the majority of studies. In studies that did not report change scores, an estimate of effect size was derived using the procedure of Smith et al. (1980), which estimates change from the pre-treatment and post-treatment group means, divided by the standard deviations reported in the original manuscript, and adjusted for test–retest correlations provided in a compendium of neuropsychological tests (Spreen and Strauss, 1998).

Weighted effect sizes, $Z$ statistics, 95% CIs, and $Q$ statistics were then calculated overall for each domain, and again within each medication group. As in Analysis 1, when the $Q$ statistic was rejected, effect sizes were combined according to the random-effects model.

Analysis 2 had a sufficient number of studies to allow for a more comprehensive examination of the influence that study characteristics might have on effect sizes and comparisons between atypical APDs. Comparisons of the dichotomous variables study blind or random assignment (controlled vs. uncontrolled), corporate sponsorship (yes vs. no), baseline medication status (typical APDs vs. unmedicated), and schizophrenia subtype classification (early phase vs. mixed) were carried out as described in Analysis 1. The variables study blind and random assignment were collapsed into a single variable due to the fact that almost every study that randomly assigned subjects to treatment was also double blind. Thus, in order to avoid the redundancy of carrying out two comparisons, studies that included at least one of these features in their design were coded as controlled and those that did not include either were coded as uncontrolled. Pearson’s $R$ correlations were carried out to examine relationships between domain effect sizes and study duration.

In addition, differences in cognitive change between medications were examined. Group differences were examined in the same manner as moderator variables, by partitioning the $Q$ statistic into a between- and within-groups component where the between-groups component reflects the difference between medication groups and the within-groups component represents an overall measure of the variability within medication groups. In cases where $Q_{BET}$ was significant, pairwise contrasts were carried out to identify specific differences between medication groups. A weighted within medication group effect size was not included in the pairwise contrasts if it was calculated under the random-effects model. The critical $a$ for the pairwise contrasts was Bonferroni corrected to control for Type I error.

Results

Analysis 1

Study demographics

Seventeen studies from 14 independent, controlled, random-assignment clinical trials were included in the analysis. The discrepancy between the number of studies and number of clinical trials is due to the fact that four studies reported on the same clinical trial of cognitive change to risperidone (Green et al., 1997; Kern et al., 1998, 1999; McGurk et al., 1997). Two studies were open label. Of the 14 independent trials, two included a clozapine arm, three included an olanzapine arm, four included a risperidone arm, two included a quetiapine arm, one included clozapine, olanzapine, and risperidone arms, one included both a risperidone and an olanzapine arm, and one included two different dose groups of quetiapine. Schizophrenia subtype classification for the 14 trials was early phase ($n=3$), general ($n=7$), and treatment refractory ($n=4$). Baseline medication status included unmedicated ($n=5$) or predominantly unmedicated ($n=1$), medicated ($n=7$), and mixed (predominantly haloperidol, $n=1$). The reported washout period for the unmedicated studies typically ranged from 2 to 7 d. After excluding four reports from the same clinical trial because of discrepancies in the reported number of enrolled subjects (Green et al., 1997; Kern et al., 1998, 1999; McGurk et al., 1997), the 13 remaining (independent) trials reported retention rates of 43–93% of enrolled patients. As expected, attrition was lower in studies with a short duration of treatment and...
retention improved to a range of 50–93% of enrolled subjects when the last observation was carried forward for analysis.

Mean trial duration was 31 wk (median = 23 wk) and ranged from 4 to 104 wk. The range of average doses used for each medication was consistent with doses recommended in the various product monographs; clozapine (410.5–521.8 mg), olanzapine (10.6–30 mg), risperidone (5.7–11.3 mg), and quetiapine (300–600 mg). The average dose used in the haloperidol control arms ranged from 4.5 to 37.9 mg.

Effect sizes

Effect sizes for one study could not be computed from the information provided by the author (Kern et al., 1998) and effect sizes for three studies were based on LOCF data. The effect size for the Global Cognitive Index was significant [effect size (ES) = 0.24, $Z = 3.67$, $p < 0.001$] indicating that atypical APDs improved overall cognitive function to a greater extent than typical APDs. The effect sizes for the Learning (ES = 0.24, $Z = 3.44$, $p < 0.001$) and Processing Speed (ES = 0.21, $Z = 3.02$, $p < 0.003$) domains were significant at the Bonferroni-corrected level. Additional improvements at the uncorrected significance level were observed for the Verbal Fluency (ES = 0.16, $Z = 2.26$, $p < 0.024$) and Motor Skill (ES = 0.21, $Z = 2.56$, $p < 0.010$) domains. The weighted mean effect sizes for the nine domains and the Global Cognitive Index are presented in Table 3. In addition, the number of subjects within each atypical medication group, summed across studies, is reported for each cognitive domain.

**Moderator variable analysis**

None of the moderator variables was significantly associated with the Global Cognitive Index (all $Q_{Bonf}$ $p$ values >0.58). Study duration and haloperidol dose used in the control arm were not significantly correlated with the Global Cognitive Index score, all Pearson’s $p$ values >0.44. Similarly, none of the moderator variables tested was associated with any domain score (all $Q_{Bonf} < 3.28$, $p > 0.070$). Effect sizes for the Cognitive Flexibility and Abstraction domain were negatively correlated with trial duration ($r = -0.70$, $p < 0.016$), however, it was apparent that this was due to an outlier (Green et al., 2002), that was significantly longer in duration (104 weeks) than the remaining studies. This correlation was not significant after removal of the Green et al. study. There was evidence that effect sizes for Processing Speed were related to the average dose used in the haloperidol control arms ($r = 0.58$, $p < 0.031$), however, this correlation

<table>
<thead>
<tr>
<th>Number of effect sizes ($k$) and number of subjects ($n$)</th>
<th>Overall weighted ES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall weighted ES</td>
</tr>
<tr>
<td></td>
<td>ES</td>
</tr>
<tr>
<td>Clozapine</td>
<td>Olanzapine</td>
</tr>
<tr>
<td>$k$</td>
<td>$n$</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Global Cognitive Index</td>
<td>3</td>
</tr>
<tr>
<td>Vigilance and Selective Attention</td>
<td>2</td>
</tr>
<tr>
<td>Working Memory</td>
<td>2</td>
</tr>
<tr>
<td>Learning</td>
<td>2</td>
</tr>
<tr>
<td>Processing Speed</td>
<td>3</td>
</tr>
<tr>
<td>Cognitive Flexibility and Abstraction</td>
<td>3</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>3</td>
</tr>
<tr>
<td>Visuospatial Processing</td>
<td>2</td>
</tr>
<tr>
<td>Motor Skill</td>
<td>1</td>
</tr>
<tr>
<td>Delayed Recall</td>
<td>3</td>
</tr>
</tbody>
</table>

ES, Effect size; CI, confidence interval.
did not remain significant when effect sizes were collapsed across groups within the studies that included multiple atypical treatment arms \((r = 0.50, p < 0.15)\).

### Analysis 2

#### Study demographics

Fifty independent groups from 41 studies met criteria for inclusion in Analysis 2. There were more groups than studies because eight studies included more than one atypical treatment arm or group. The schizophrenia subtype classification included early phase \((n = 5)\), general \((n = 18)\), and treatment-refractory \((n = 18)\) patients. Baseline medication status included unmedicated \((n = 11)\), medicated \((n = 22)\), mixed \((\text{unmedicated} / \text{typicals} = 1; \text{typicals} / \text{atypicals} = 4)\), and unknown \((n = 3)\). Washout periods for the unmedicated studies typically ranged from 1 to 7 d although one study included only neuroleptic-naive subjects. Eighteen studies either randomly assigned subjects to treatment or were double blind. Eighteen studies received at least partial funding support from a pharmaceutical company. Among the studies that were not included in Analysis 1, the percentage of subjects completing the trials ranged from 45% to 100%. As expected the average percentage was high, 82%, possibly reflecting the tendency for less controlled studies to infrequently report the number of subjects initially screened or enrolled in a study. Mean trial duration was 25 wk (median = 14 wk) and ranged from 1.5 wk to 3 yr. The mean and range (in parentheses) of doses under double-blind (DB) conditions tended to be lower than the open-label (OL) doses in studies of clozapine \([\text{DB} = 454.3 (410.5–498), \text{OL} = 478.4 (200–719)]\), and quetiapine \([\text{DB} = 424.5 (300–600), \text{OL} = 529.1 (319.3–750)]\), whereas the reverse was true for olanzapine \([\text{DB} = 16.9 (10–30), \text{OL} = 13.8 (11–19.9)]\), and risperidone \([\text{DB} = 7.7 (5.7–11.3), \text{OL} = 5.5 (2.2–8.9)]\).

#### Effect sizes

The results for Analysis 2 are shown in Table 4. The Global Cognitive Index for all atypical treatments combined was significantly greater than zero \((ES = 0.36, Z = 8.87, p < 0.001)\). All cognitive domains demonstrated significant improvement on atypical APD medications at the Bonferroni-corrected significance level. The weighted effect sizes for the nine domains ranged from 0.17 to 0.46. The weighted effect sizes for the Vigilance and Selective Attention, Learning, and Delayed Recall domains were calculated under the random-effects model due to the presence of significant heterogeneity (all \(\chi^2 p\) values < 0.010). Inspection
Meta-analysis of cognitive change

Figure 1. Neuropsychological change to atypical antipsychotic drugs: controlled (■) vs. uncontrolled studies (□). * Indicates significant differences between controlled and uncontrolled trials \((p < 0.05)\). For abbreviations see Table 1.

of the distribution of effect sizes within the Learning domain revealed an outlier \((ES = 1.22)\) that was significantly greater than the range of the remaining effect sizes \((ES = -0.13 - 0.84)\). Removal of this outlier significantly reduced the variation within the Learning domain \((\chi^2_{df=14}=28.95, p < 0.022)\) and the mean effect size remained significant \((ES = 0.43, Z = 9.94, p < 0.001)\). This outlier is not included in the moderator or comparisons between treatment groups’ analyses below.

Moderator variables

Global Cognitive Index effect sizes from uncontrolled studies \((ES = 0.43, Z = 6.75, p < 0.001)\) were marginally larger than those from controlled studies \((ES = 0.32, Z = 6.03, p < 0.001)\), however, this difference was not significant, \(Q_{BET} = 1.95, p < 0.164\). The moderator variable, control, was significantly associated with Verbal Fluency effect sizes \((Q_{BET} = 8.19, p < 0.005; Q_W = 32.39, p < 0.595; R^2 = 0.18)\) and Processing Speed effect sizes \((Q_{BET} = 6.82, p < 0.009; Q_W = 47.73, p < 0.252; R^2 = 0.11)\). Verbal Fluency effect sizes calculated from random assignment or double-blind studies were significantly less than those obtained from open-label, uncontrolled studies \((ES = 0.21 vs. 0.45)\). Similarly, effect sizes for Processing Speed were also larger in the uncontrolled relative to controlled studies \((ES = 0.50 vs. 0.30)\). The weighted mean and 95% CI for each domain and the Global Cognitive Index for controlled and uncontrolled studies are displayed in Figure 1. Within the controlled studies, the weighted mean effect size for each domain remained significant after Bonferroni correction, however, the weighted mean effect sizes for the Vigilance and Selective Attention, Cognitive Flexibility and Abstraction, and Visuospatial Skill domains calculated from uncontrolled studies were not. The moderator variable baseline medication status was significantly associated with Delayed Recall domain effect sizes \((Q_{BET} = 5.98, p < 0.015; Q_W = 26.29, p < 0.0240; R^2 = 0.14)\). Studies that included an unmedicated baseline produced smaller Delayed Recall effect sizes than those that tested subjects while they were receiving typical APDs at baseline \((ES = 0.21 vs. 0.54)\). The moderator variables diagnosis, corporate sponsorship, and schizophrenia subtype were not significantly associated with the Global Cognitive Index score or any domain. Trial duration was not correlated with the Global Cognitive Index or any domain effect size.

Comparison of atypical antipsychotic drugs

The \(Q_{BET}\) statistic revealed significant group differences within the Vigilance and Selective Attention domain \((Q_{BET} = 22.53, p < 0.001; Q_W = 26.52, p < 0.491; R^2 = 0.46)\) and the Verbal Fluency domain \((Q_{BET} = 15.47, p < 0.002; Q_W = 25.18, p < 0.912; R^2 = 0.32)\). Within the Vigilance and Selective Attention domain, follow-up contrasts identified a significant advantage for quetiapine, relative to clozapine \((\chi^2_{df=1} = 11.35, p < 0.001)\) or risperidone \((\chi^2_{df=1} = 15.47, p < 0.001)\), and a significant advantage for olanzapine, relative to risperidone \((\chi^2_{df=1} = 10.92, p < 0.001)\) at the Bonferroni-corrected level \([p = (0.05/6) = 0.008]\). Additional advantages for quetiapine, relative to olanzapine \((\chi^2_{df=1} = 4.19, p < 0.041)\), and olanzapine, relative to clozapine \((\chi^2_{df=1} = 5.41, p < 0.021)\), were observed at the uncorrected significance level. Pairwise contrasts within the Verbal Fluency domain indicated that quetiapine improved performance to a greater extent than risperidone \((\chi^2_{df=1} = 11.09, p < 0.001)\), and clozapine improved verbal fluency to a greater extent than risperidone \((\chi^2_{df=1} = 9.19, p < 0.003)\) after Bonferroni correction. Additional advantages for quetiapine, compared to olanzapine \((\chi^2_{df=1} = 6.26, p < 0.013)\), and clozapine, compared to olanzapine \((\chi^2_{df=1} = 3.86, p < 0.050)\), were observed at the uncorrected significance level. The Verbal Fluency pairwise contrasts were repeated after exclusion of the uncontrolled studies since this moderator variable was associated with verbal fluency effect sizes. After excluding uncontrolled studies, the quetiapine vs. risperidone and quetiapine vs. olanzapine contrasts were significant at the Bonferroni-corrected significance level \((\chi^2_{df=1} = 10.16, p < 0.002)\) and \(\chi^2_{df=1} = 7.54, p < 0.007\) respectively, but the clozapine vs. risperidone contrast was not \((\chi^2_{df=1} = 4.50, p < 0.034)\).
Within-group effect sizes

The within-group effect sizes for each medication are presented in Table 4. Because the moderator control was significant for Verbal Fluency and Processing Speed, the within-group effect sizes for these two domains were recalculated after removing uncontrolled studies. After excluding the uncontrolled studies, the Verbal Fluency effect sizes for clozapine (ES = 0.41, Z = 2.87, p < 0.005), and quetiapine (ES = 0.68, Z = 3.92, p < 0.001) remained significant. However, the Verbal Fluency effect size for olanzapine (ES = 0.17, Z = 2.54, p < 0.012) and the Processing Speed effect sizes for clozapine (ES = 0.28, Z = 0.99, p < 0.322), and risperidone (ES = 0.19, Z = 2.10, p < 0.036) did not. The results for quetiapine should be interpreted cautiously given that the effect sizes for several domains included relatively few studies and, in the case of visuospatial processing, were based on a single study.

Discussion

The findings from the current set of meta-analyses indicate that atypical APDs improve overall cognitive function in schizophrenia and performance in a number of cognitive domains. The results obtained from Analysis 1 of 14 controlled, random-assignment trials indicates that atypical APDs are superior to typical APDs, haloperidol in particular, at improving overall cognitive function. This finding is consistent with an earlier meta-analysis of three randomized, controlled trials that identified improvement in overall cognitive function with atypical APDs. In contrast to the earlier meta-analysis that was based upon a small number of clinical trials conducted prior to 1999, the greater number of studies in the current meta-analysis allowed for a closer examination of the improvements. After Bonferroni correction, improvements were identified in learning and processing speed. Additional improvements in verbal fluency and motor skill were detected, although these improvements failed to reach Bonferroni-corrected significance levels.

The inclusion of investigations with single treatment arms and uncontrolled designs in Analysis 2 further supports the benefits of atypical APD treatments and indicates improvements occur in a wide array of cognitive functions. The effect sizes for domains ranged from 0.17 to 0.46 and are remarkably consistent with Harvey and Keefe’s (2001) earlier review of 20 studies. For example, Harvey and Keefe (2001) identified improvements, in terms of Cohen’s d, of 0.39 and 0.18 for vigilance and executive functions respectively. The results reported here for Vigilance and Selective Attention and Cognitive Flexibility and Abstraction were 0.34 and 0.17.

A primary advantage of the meta-analytical strategy involves the ability to analyse moderator variables. There was no compelling evidence that moderator variables influenced effect sizes among the set of randomized, controlled trials. However, a trend emerged for a positive correlation between haloperidol dose and the degree to which the patients treated with atypical APDs outperformed haloperidol-treated patients on processing speed tasks. Although this association failed to reach statistical significance, it suggests that some of the advantages of atypical APDs relates to an avoidance of the deleterious effects of high doses of haloperidol. Alternatively, one might speculate that this association reflects symptom severity, with the most severe patients requiring the highest treatment doses, and the most severe subjects showing the greatest relative advantage of atypical APDs.

Examination of moderator variables in Analysis 2 revealed that studies that failed to randomly assign subjects to treatment or utilized open-label designs produced larger verbal fluency and processing speed effect sizes than studies that included either of these features in their design. However, caution is warranted when interpreting these findings, particularly with respect to verbal fluency, because the larger number of clozapine studies within the group of open-label studies may have skewed the results. Clozapine, in contrast to olanzapine and risperidone, significantly improves verbal fluency in both open-label and double-blind studies and it is possible that the higher number of open-label clozapine studies may have inflated the mean effect size. The Global Cognitive Index was not significantly different between controlled and uncontrolled studies suggesting that study methodology does not systematically bias all results. Rather, the effects of study design appear to increase the variability of effects across studies as evidenced by the fact that, within any given domain, uncontrolled studies yielded a broader range of effect sizes than the controlled studies.

Pairwise contrasts between atypical APDs indicated that no medication appeared superior or inferior to the other medications in overall cognitive function, but several differences emerged in two domains, Vigilance and Selective Attention, and Verbal Fluency. The findings should be considered preliminary until more large-scale, controlled comparisons between atypical APDs are carried out, particularly with clozapine and quetiapine. However, the results are generally consistent with predictions derived from the assumption...
that lower dopamine D2 receptor affinity and increased serotonergic effects may be related to cognitive benefits from novel agents. In contrast the results are not entirely consistent with the longstanding assumption that the inherent anticholinergic properties of some APDs might limit gains in memory and attention (McGurk and Powchick, 2000). Risperidone, which has the highest affinity for D2 receptors among the atypical agents (Schotte et al., 1996; Seeman, 2002), showed the least beneficial profile on measures of Vigilance and Selective Attention and Verbal Fluency, being outperformed by quetiapine and olanzapine on Vigilance and Selective Attention, and quetiapine and clozapine on Verbal Fluency. The differences were quite robust ranging from 0.3 to 0.5 standard deviations and, for Verbal Fluency, remained significant even when the analysis was restricted to controlled studies. Clozapine, which may be more cholinomimetic than anticholinergic (Olianas et al., 1999; Zorn et al., 1994), did not significantly improve Vigilance and Selective Attention and it resulted in less improvement than quetiapine on this domain. Moreover, although clozapine significantly improved Delayed Recall, improvement in this domain was markedly less than that observed in the olanzapine and risperidone groups. However, despite the presumption of significant inherent anticholinergic activity, olanzapine did not conform to this model. Olanzapine treatment produced medium to large gains on tests of vigilance and selective attention and delayed recall. It thus appears that, at least at clinically relevant dosages, olanzapine does not appear to behave like an anticholinergic agent. These conclusions are consistent with the absence of further cognitive impairment observed in patients with Alzheimer’s disease treated with very low doses of olanzapine (Kennedy et al., 2001; Street et al., 2000) and the lower incidence of cholinergic-related side-effects and serum anticholinergic levels observed with olanzapine relative to clozapine (Chengappa et al., 2000; Eschweiler et al., 2002).

The moderator analysis is an effective method for detecting systematic variability between different studies of cognitive change to novel treatments, but it does not allow an assessment of more systemic challenges to the validity of the cognitive benefits reported from atypical APDs relative to typical APDs or to the validity of differential benefits within the atypical APD class. One factor especially germane to the current review is the adjunctive use of anticholinergic medications. In studies with a typical APD control arm, emergent EPS require adjunctive anticholinergic medication that may interfere with cognitive skills, particularly attention and memory. It is notable in this regard that delayed recall scores showed the largest improvements with atypical APDs in studies in which subjects were originally assessed while on a typical APD.

A second artifact relates to the possibility of practice effects that could occur on neuropsychological measures that are repeatedly administered to the same subject. In atypical APD vs. typical APD studies, practice effects would be expected in both treatment arms, thus, a relative advantage of atypical APDs probably would not be related to practice effects alone. However, this inference relies on the unsupported assumption that there will be no interaction between treatment and practice (Carpenter and Gold, 2002). To the contrary, emerging evidence suggests that typical APD treatments may have subtle, detrimental effects on cognition that may limit the benefit of repeated exposure to the same materials (Blyler and Gold, 2000). For example, normalization of procedural learning following a change from atypical APDs to clozapine suggest that some improvements in cognitive function may relate to a release from impairment caused by the typical APD (Purdon et al., 2002). Similar demonstrations of a preservation of procedural learning with olanzapine and clozapine compared to the apparent loss of procedural learning induced by haloperidol, and perhaps risperidone (Bedard et al., 1996, 2000; Purdon et al., 2003; Stevens et al., 2002) support the view that some of the improvements with atypical APDs might result from an avoidance of deleterious effects on learning associated with typical APDs. While typical APDs may limit practice effects, the improvements on atypical APDs are unlikely to be entirely explained by practice effects. The percentage of patients demonstrating improvements at or greater than half a standard deviation, which ranges from 40% to 75%, in recent double-blind, controlled trials (Bilder et al., 2002; Harvey et al., 2003; Velligan et al., 2003) exceeds what one would expect from typical practice effects. Moreover, the differences between atypicals on verbal fluency and attention in the current study can not be accounted for by practice effects.

As with any meta-analysis, publication bias, especially among studies sponsored by pharmaceutical companies, poses a threat to the validity of the findings. Corporate sponsorship plays a significant role in the dissemination of results and although there was no evidence that sponsored trials reported larger effect sizes, it remains possible that a number of sponsored, unpublished negative trials exist. Selective reporting of results within published papers can also pose a threat. However, almost all of the trials included in
the current review examined multiple dependent measures simultaneously and reported all the results within a single article, although there were exceptions (see Green et al., 1997; Kern et al., 1998, 1999; McGurk et al., 1997 for exceptions). Thus, while we cannot rule out the existence of unpublished negative findings, within the published studies analysed here, it seems unlikely that a systematic positive reporting bias exists.

The improvements in cognitive performance with atypical APDs are in general encouraging, especially when the potential implications for socio-vocational re-integration are considered. The gains observed in learning may be particularly relevant as this cognitive skill has been linked to three major dimensions of outcome including community/daily activities, social problem solving/instrumental skills, and psychosocial skill acquisition (Green, 1996; Green et al., 2000). However, it is prudent to conclude this discussion with emphasis on the relatively small magnitude of the observed changes. Schizophrenia patients typically score more than a standard deviation below healthy controls on many of the neuropsychological tests reviewed here (Heinrichs and Zakzanis, 1998). As a class, atypical APDs improve overall cognitive function but the improvement is typically in the range of 0.20–0.40 standard deviations. It is highly unlikely that the gains will be sufficient to return patients to the vocational level predicted from their individual pre-morbid status. However, the medication-specific effects of particular atypical APDs on particular cognitive domains could be relevant to the design of individual treatment plans that take into account the patient’s pre-morbid intellect, unique profile of cognitive impairment, prior vocational achievements, and long-term socio-vocational aspirations.

Acknowledgements
The authors thank Dr Sohee Park and Dr Bahr Weiss for their helpful comments on an earlier draft of this manuscript. The authors also thank all the researchers who kindly provided additional study data upon request.

Statement of Interest
Mr Neil D. Woodward, M.A. and Dr David H. Zald, Ph.D. have no affiliation or financial or other relationships with any organization with a financial interest in the subject matter or material discussed in the manuscript. Dr Scot E. Purdon, Ph.D. has received grant/research support, served on an advisory board and/or as a speaker for Eli Lilly & Co. Dr Herbert Y. Meltzer, M.D. has served as a consultant, board member and/or speaker for Janssen, Novartis, and Pfizer and has received grant/research support from AstraZeneca, Eli Lilly & Co., Janssen, Novartis, and Pfizer. No financial support for this manuscript was received from any manufacturer of the pharmaceutical products discussed in this review.

References


Carpenter WT, Gold JM (2002). Another view of therapy for cognition in schizophrenia. Biological Psychiatry 51, 969–971.


Chua L, Chong SA, Pang E, Ng VP, Chan YH (2001). The effect of risperidone on cognitive functioning in a sample
of Asian patients with schizophrenia in Singapore.


Harvey PD, Siu CO, Romano S (2004). Randomized, controlled, double-blind, multicenter comparison of the cognitive effects of ziprasidone versus olanzapine in acutely ill inpatients with schizophrenia or schizoaffective disorder. Psychopharmacology (Berlin) 172, 324–332.


5-HT(2A) and D(2) receptor blockade increases cortical DA release via 5-HT(1A) receptor activation: a possible mechanism of atypical antipsychotic-induced cortical dopamine release. *Journal of Neurochemistry* 76, 1521–1531.


