

Sex Difference in Cognitive Response to Antipsychotic Treatment in First Episode Schizophrenia

Leah H Rubin^{*1}, Gretchen L Haas², Matcheri S Keshavan^{2,3}, John A Sweeney¹ and Pauline M Maki¹

¹Department of Psychiatry and Psychology, Center for Cognitive Medicine, University of Illinois at Chicago, Chicago, IL, USA; ²Department of Psychiatry, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; ³Department of Psychiatry, Wayne State University School of Medicine, Detroit, MI, USA

We extend the investigation of cognitive sex differences in schizophrenia in a novel way by grouping cognitive tests according to the direction of the typical sex difference, an approach used in studies of hormonal effects on behavior in other clinical conditions. Additionally, we explore how performance on these 'male' and 'female' tests changed following antipsychotic treatment. Seventy patients with a first hospitalization for schizophrenia or schizophreniform disorder completed cognitive tests before antipsychotic treatment and approximately 5 weeks after treatment. Thirty-nine healthy comparison subjects completed the tests at similar intervals. Primary outcome variables were composite scores on cognitive tests that in normative studies favor males over females ('male' tests) or favor females over males ('female tests'). Overall, patients performed more poorly than healthy individuals. The expected pattern of sex differences was found on the composite test scores, with an advantage for females on 'female' tests and an advantage for males on 'male' tests. Female patients showed a greater improvement on 'female' tests and a decrease in performance on 'male' tests following treatment. Although male patients did not perform significantly better after treatment on 'female' tests, they did improve on non-motor 'male' tests of visuospatial skills. Future studies of the neurocognitive effects of antipsychotic treatment may need to take potential sex differences in cognitive response into account.

Neuropsychopharmacology (2008) **33**, 290–297; doi:10.1038/sj.npp.1301395; published online 28 March 2007

Keywords: schizophrenia; sex differences; first episode; treatment; cognition

INTRODUCTION

There is compelling evidence that sex modulates the clinical presentation and course of schizophrenia. The course of the disease is more benign in women than in men; females on average have a later age of onset, shorter and less frequent acute psychotic episodes, less severe negative symptoms, better premorbid functioning, and a better treatment response to antipsychotic medication compared to males (Grigoriadis and Seeman, 2002). However, evidence remains equivocal regarding sex differences in neurocognitive deficits and in the response of these deficits to antipsychotic treatment. Some neuropsychological studies find men to be more impaired than women (Goldstein *et al*, 1998; Haas *et al*, 1991; Kopala *et al*, 1989; Seidman *et al*, 1997); others report women to be more impaired than men (Goldberg *et al*, 1995; Lewine *et al*, 1996; Perlick *et al*, 1992); and some have not observed differences between the sexes (Albus *et al*, 1997; Hoff *et al*, 1998). For example, Goldstein *et al*

(1998) reported that female patients outperformed male patients on tests of attention, verbal memory, and executive functioning. By contrast, Goldberg *et al* (1995) generally found similar performances between the sexes on approximately 100 neuropsychological measures, and the few significant differences favored men over women. Gur *et al* (2001) similarly found no overall differences in neurocognitive profiles in a sample of clinically stable patients with schizophrenia. However, similar to reports in healthy subjects, female patients performed better on a verbal memory task and were more impaired on a visuospatial task compared to male patients (Gur *et al*, 2001). This finding has been recently replicated in another sample of clinically stable patients with schizophrenia (Halari *et al*, 2006). Common explanations for discrepant findings regarding sex differences in cognition include variations in illness severity (acute, chronic), inadequate sample size, sources of sampling bias, and non-balanced samples (more men than women).

Sex differences in cognitive abilities are well documented. For instance, women excel at tasks involving fine motor skills and verbal fluency, whereas men excel at visuospatial abilities (Kimura and Hampson, 1993). One framework for understanding sex differences in cognitive functioning is the behavioral endocrinological approach that holds that

*Correspondence: LH Rubin, Center for Cognitive Medicine, Neuropsychiatric Institute, University of Illinois at Chicago, 912 S. Wood Street MC913, Chicago, IL 60612, USA, Tel: +1 312 355 5652, Fax: +1 312 413 7856, E-mail: lrubin@psych.uic.edu

Received 24 May 2006; revised and accepted 15 February 2007

sex differences often reflect the permanent effects of prenatal sex steroid hormones on brain structure and function (organizational effects) as well as the physiologic actions of circulating sex steroid hormones on brain structure and function in later life (activational effects). In this context, an approach used in the field of behavioral endocrinology to study cognitive sex differences is to group tasks according to sex differences in normative samples, using a 'male' score comprised of tests on which males outperform females and a 'female' score for tests on which females on average perform at a higher level compared to males. This approach contrasts the standard approach of analyzing sex differences in cognitive domains (eg, memory, executive function) comprised of varying tests regardless of normative sex differences in the ability to perform them. The behavioral endocrinological approach has been used in studies examining influences of sex steroid hormones on cognitive performance. For example, a study examining the possible influence of estrogen on cognitive abilities in healthy young women demonstrated within-subject enhancements on a 'female' composite score during the high-estrogen phase of the menstrual cycle compared to the low-estrogen phase of the cycle (Hampson, 1990). Conversely, a study of male-to-female trans-sexuals found a decrease in performance on a 'male' composite score with estrogen supplementation and an increase on that composite score in female-to-male trans-sexuals following testosterone administration (Slabbekoorn *et al*, 1999). To date, this approach of grouping tests according to the direction of sex differences has not been used in studies of sex differences in cognitive functioning in untreated first-episode schizophrenia, or to our knowledge in the study of treatment effects on cognition in this disorder.

In the present study, we sought to investigate sex differences in cognitive response to medication in first-episode patients with schizophrenia by examining performance on 'male' and 'female' neurocognitive composite scores. Specifically, we examined performance before and after antipsychotic treatment in a sample of antipsychotic-naïve first-episode schizophrenia patients participating in the University of Pittsburgh First-Episode Project. We hypothesized that (a) women with schizophrenia would show a post-treatment selective improvement in cognitive domains in which females typically show an advantage, and (b) this response would be greater than any change observed in male patients on cognitive tests that favor males or that favor females. A finding of greater sex-specific cognitive enhancement in female compared to male patients would suggest effects beyond a generalized 'normalization' of sex differences post-treatment.

METHODS

Participants

The sample included 70 patients (44 males, 26 females) and 39 healthy individuals (23 males, 16 females) recruited into the University of Pittsburgh First-Episode Project, a prospective study of the early course of schizophrenia. Diagnoses of these antipsychotic naïve patients were made according to The American Psychiatric Association (APA) DSM IV, 1994 using the Structured Clinical Interview for

DSM diagnoses (SCID) (Spitzer *et al*, 1992) and consensus diagnostic review meetings examining all available clinical information. Healthy individuals were matched to patients on age, sex, and parental socioeconomic status (see Table 1). Healthy subjects in the project were recruited from the community via bulletin board and newspaper advertisements and had no Axis I diagnoses based on the SCID. Inclusion criteria included (1) age between 14 and 45 and (2) English as first language. Exclusion criteria included: (1) history of systemic or neurological disease, (2) earlier treatment with electroconvulsive therapy, (3) history of head trauma with loss of consciousness, (4) history of substance dependence or substance abuse within 3 months of study enrollment, and (5) history of medications known to affect neurobehavioral functions within 1 month of study entry. All female patients and healthy individuals were premenopausal. Twenty-five female patients were not receiving oral contraceptives, and for one patient this information could not be confidently determined. Seven female healthy subjects were on oral contraceptives, three were not on oral contraceptives, and for six women this information was not available.

Procedures

All participants gave oral and written consent, and the University of Pittsburgh Institutional Review Board (IRB) approved the study. A neuropsychological test battery was administered to patients no more than 7 days before treatment initiation, and follow-up testing was conducted, on average, 39 days (SD = 16) after treatment began. Within 1 week of each neuropsychological testing, the following clinical scales were administered by individuals with no knowledge of neuropsychological test findings: Brief Psychiatric Rating Scale (BPRS) (Overall and Gorman, 1962), the schedule for assessment of positive symptoms (SAPS) (Andreasen, 1994a), the schedule for assessment of negative symptoms (SANS) (Andreasen, 1994b), the 24-item Hamilton Depression Rating Scale (HAM-D) (Hamilton, 1960), and the global assessment of functioning (GAF) (Endicott *et al*, 1976). Following initial testing, patients were treated with either typical antipsychotics, including

Table 1 Demographics of Schizophrenia Patients and Healthy Subjects at Baseline as a Function of Sex

	Healthy subjects		Patients	
	Men	Women	Men	Women
N	26	16	44	26
Age	23.14 (4.95)	25.64 (5.79)	23.84 (6.68)	26.69 (8.32)
IQ	100.48 (6.34)	105.38 (11.03)	97.98 (11.54)	96.80 (10.86)
Parental SES	2.48 (0.99)	2.18 (0.55)	2.41 (1.22)	2.68 (1.05)
Race (%)				
Caucasian	70	75	73	42
African-American	30	25	21	54
Other	0	0	6	4

haloperidol ($n = 20$, median dosage = 3 mg, range 1–15 mg), fluphenazine decanote ($n = 1$, dosage = 1 mg), perphenazine ($n = 3$, median dosage = 6 mg, range 4–8 mg), loxapine ($n = 1$, dosage = 20 mg), or atypical medications including risperidone ($n = 36$, median dosage = 3 mg, range 1–12 mg), olanzapine ($n = 1$, dosage = 20 mg), quetiapine ($n = 1$, dosage = 25 mg), or unknown antipsychotic medication (risperidone or haloperidol) owing to randomized clinical trial participation ($n = 7$, median dosage = 4 mg, range 4–6 mg). Patients were treated by clinician choice except for the clinical trial patients.

Measures

The neuropsychological tests selected for this study were selected from a more extensive test battery (Hill *et al*, 2004) based on previous research indicating that the tests demonstrate sex differences in the general population. The following tests were selected and grouped according to existing empirical data demonstrating reliable sex differences in normative samples.

Female tests. California Verbal Learning Test (CVLT) (Delis *et al*, 1987): a measure of immediate recall, delayed recall, and delayed recognition. The outcome measure was a composite score comprised of the total number of words recalled across the five learning trials, the total number of words recalled on the short-delay free recall, and the total number of words recalled on the long-delay free recall.

Digit Symbol Subtest of the Wechsler Adult Intelligence Scale (DSYM) (Wechsler, 1981): a measure of perceptual-motor skills. The outcome measure was the total number of correct items.

Grooved Pegboard Test (GPEG) (Reitan and Wolfson, 1985): a measure of fine motor skills. The outcome measure was the average time to complete the task for the dominant and non-dominant hands.

Controlled Oral Word Association Test (COWAT) (Thurstone and Thurstone, 1938): a measure of letter verbal fluency. The outcome measure was the total number of words generated to letter cues across three 60-sec trials (F, A, S).

Previous studies report sex differences in favor of females on each of these tests: each of the specific CVLT outcome measures comprising the composite CVLT score (Kramer *et al*, 1988, 1997), DSYM (McCurry *et al*, 2001; Snow and Weinstock, 1990; Mann *et al*, 1990), GPEG (McCurry *et al*, 2001; Ruff and Parker, 1993; Schmidt *et al*, 2000, and COWAT (Burton *et al*, 2005; Kimura, 1992; Kimura and Hampson, 1993).

Male tests. Finger Tapping Test (Reitan and Wolfson, 1985): a measure of motor speed. The outcome measure was the average number of total presses across the 10 trials per hand for the right and left hand (FTAP).

Benton Judgment of Lines Orientation Test (JOLO) (Benton *et al*, 1983): a measure of visuospatial abilities. The outcome measure was the total number of correct items (30 trials).

Previous studies report sex differences in favor of males on these tests: FTAP (Ruff and Parker, 1993; Chavez *et al*,

1983; Morrison *et al*, 1979; Shimoyama *et al*, 1990) and JOLO (Rahman and Wilson, 2003; Riva and Benton, 1993).

Psychometrics. Internal consistency using Cronbach's α was computed for the overall sample at baseline for the four 'female' tests (CVLT composite, DSYM, GPEG, COWAT) and for the two 'male' tests (Finger Tapping, JOLO) and yielded estimates of 0.64 and 0.51, respectively. For the overall sample, test-retest reliabilities for 'female' and 'male' tests were 0.82, and 0.80, respectively. There were no group differences on internal or test-retest reliability on 'female' or 'male' tests at baseline or following treatment.

Data Analyses

Less than 1% of the data was missing. Missing values were imputed using the SPSS missing value analysis regression technique with random residuals. Predictors used to estimate missing values for individual test scores included neuropsychological test scores at baseline and at follow-up (continuous predictors), sex, group, and race. Cognitive test scores were transformed into standardized z-scores using data obtained from the healthy group (males and females combined) at baseline testing. The resulting z-scores were first averaged into within-test composite scores. For example, the CVLT composite was calculated by averaging the z-score for total words recalled across the five trials, short-delay free recall, and long-delay free recall. Next, these within-test composite scores were averaged with z-scores from other sexually dimorphic tests to create two separate composites representing 'female' and 'male' advantaged cognitive abilities.

Primary analyses. A multi-factor ANOVA was used to test the primary hypotheses. The grouping variables were Sex (men, women) and Group (healthy, patients), and the within-subject variables were Time (baseline, 5 weeks) and Test ('female', 'male'). Of particular interest for all hypotheses was the four-way Group \times Sex \times Time \times Test interaction followed by specific simple interactions to test specific hypotheses. For example, the follow-up test of the three-way Group \times Sex \times Test interaction at baseline tested the hypothesis that sex differences would be greater in controls at baseline compared with patients. All follow-up tests were computed using the appropriate error term from the primary mixed factors analysis. Greenhouse-Geisser corrected p -values were used to control for family-wise error.

RESULTS

Table 2 shows mean z-scores (with standard error) for each neuropsychological measure and the 'female' and 'male' overall composite scores for the two groups at the two test sessions. Healthy subjects performed better than patients across all domains, $F(1, 105) = 39.04$, $p < 0.001$. The Sex \times Test interaction was also significant, validating the designation of the various tests as 'female' or 'male', $F(1, 105) = 21.07$, $p < 0.001$. Figure 1 shows the change in cognitive performance for male and female patients and healthy subjects on 'female' (a) and 'male' (b) tests from baseline to post-test. The test of the primary hypo-

Table 2 Mean and Standard Error of z-Scores as a Function of Study Task, Test Type, Sex, and Time for Group

Test	Patients				Healthy subjects			
	Men (n = 44)		Women (n = 26)		Men (n = 23)		Women (n = 16)	
	Baseline	Post-treat	Baseline	Post-treat	Baseline	5 weeks	Baseline	5 weeks
<i>Female</i>								
CVLT composite	−0.86 (0.21)	−0.98 (0.23)	−0.87 (0.26)	−0.50 (0.29)	−0.08 (0.21)	0.23 (0.18)	0.11 (0.19)	0.46 (0.16)
Digit symbol ^a	−1.76 (0.16)	−1.78 (0.19)	−1.65 (0.17)	−1.31 (0.20)	−0.48 (0.13)	−0.17 (0.16)	0.69 (0.26)	0.66 (0.35)
Grooved pegboard composite ^a	−0.86 (0.20)	−1.08 (0.20)	−0.78 (0.27)	−0.84 (0.28)	−0.33 (0.23)	−0.10 (0.16)	0.48 (0.15)	0.55 (0.16)
COWAT	−0.74 (0.21)	−0.57 (0.24)	−1.30 (0.22)	−0.73 (0.27)	0.11 (0.22)	0.18 (0.23)	−0.15 (0.24)	0.09 (0.26)
Female composite	−1.06 (0.11)	−1.11 (0.13)	−1.15 (0.14)	−0.84 (0.17)	−0.20 (0.15)	0.04 (0.18)	0.28 (0.18)	0.35 (0.29)
Female non-motor composite	−1.12 (0.14)	−1.11 (0.18)	−1.27 (0.13)	−0.85 (0.17)	−0.15 (0.10)	0.08 (0.10)	0.22 (0.14)	0.40 (0.17)
<i>Male</i>								
Judgment of lines ^b	−1.34 (0.36)	−0.65 (0.35)	−2.05 (0.40)	−2.52 (0.54)	0.29 (0.15)	0.30 (0.19)	−0.41 (0.30)	−0.49 (0.29)
Finger tapping composite ^b	0.28 (0.21)	−0.29 (0.22)	−0.64 (0.28)	−0.84 (0.24)	0.27 (0.22)	0.40 (0.21)	−0.39 (0.20)	−0.13 (0.23)
Male composite	−0.53 (0.21)	−0.47 (0.21)	−1.35 (0.26)	−1.68 (0.27)	0.28 (0.28)	0.35 (0.29)	−0.40 (0.33)	−0.31 (0.34)

^aSignificant sex difference in favor of females in healthy subjects at baseline.

^bSignificant sex difference in favor of males in healthy subjects at baseline. Also, recall that z-scores were calculated by creating within-test z-scores. This is important because the mean of the baseline z-scores for a multi-variable test such as the CVLT is not 0 for the healthy subjects as one would expect if all scores were converted to z-scores based on the baseline scores for the total healthy sample.

theses—the four-way Group \times Sex \times Time \times Test interaction—was significant, $F(1, 105) = 4.75$, $p = 0.04$. This indicated that the magnitude of change in performance after treatment depended on the combined influence of treatment, sex, and whether tests favor men or women in the general population (and in this sample). The Sex \times Test \times Time interaction was significant for patients, but not for healthy subjects, $F(1, 105) = 6.10$, $p < 0.05$, and $F(1, 105) = 0.34$, NS, respectively, indicating that the significant change over time in test performance of men and women on ‘male’ and ‘female’ tests was restricted to the patient group. Restricting a follow-up analysis to patients of each sex, the Test \times Time interaction was significant for female patients, but not male patients, $F(1, 105) = 11.60$, $p < 0.05$, and $F(1, 105) = 0.6$, NS. As expected, female patients showed significant improvement on the ‘female’ tests and significant decline in performance on the ‘male’ tests, $F(1, 105) = 5.32$, $p < 0.05$, and $F(1, 105) = 6.30$, $p < 0.05$, respectively. Overall these results indicate that female patients showed a specific and differential improvement on ‘female’ tests and decline on ‘male’ tests following treatment, whereas male patients did not show a change in test performance on ‘male’ or ‘female’ tests after treatment. As predicted, the magnitude of the change in performance in female patients on ‘female’ tests (0.31 SD) was greater than the magnitude of improvement in male patients on cognitive tests that favor males (0.06 SD) or that favor females (0.05 SD).

The standardized mean difference (SMD) method was used to calculate effect sizes for the ‘female’ and ‘male’ tests within female patients and in healthy subjects to evaluate the magnitude of the treatment effect. Effect sizes were calculated within female patients and healthy subjects by subtracting their baseline score on the ‘female’ and ‘male’ tests by their retest scores. On ‘female’ tests, the effect size

was nearly twice as large for female patients compared to healthy female subjects, $ES = 0.31$ and $ES = 0.16$, respectively. Similarly on ‘male’ tests, there was a larger effect size in female patients compared to healthy female subjects, $ES = -0.33$ and $ES = -0.09$, respectively.

A *post-hoc* test of the primary hypothesis was conducted to exclude tests that predominantly required motor control, specifically the grooved pegboard and finger tapping test. This analysis was done in light of possible adverse effects of antipsychotic treatment on motor skills (ie, adverse effects on extrapyramidal systems) that might counteract a beneficial effect on higher cognitive functions in our composite score analyses. Again, the four-way Group \times Sex \times Time \times Test interaction was significant, $F(1, 105) = 6.23$, $p = 0.01$. The simple Sex \times Test \times Time interaction again was significant for patients, but not for healthy subjects, $F(1, 105) = 9.19$, $p < 0.05$, and $F(1, 105) = 0.47$, NS, respectively. Restricting the analyses to patients, the Test \times Time interaction was significant for both female and male patients, $F(1, 105) = 9.24$, $p < 0.05$, and $F(1, 105) = 9.16$, $p < 0.05$, respectively. Female patients showed significant improvement on the ‘female’ tests and significant decline in performance on the ‘male’ tests, $F(1, 105) = 11.77$, $p < 0.05$, and $F(1, 105) = 5.07$, $p < 0.05$, respectively. Conversely, male patients showed significant improvement on the ‘male’ composite (now consisting only of JOLO) and no improvement on the ‘female’ composite, $F(1, 105) = 18.57$, $p < 0.05$, and $F(1, 105) < 1.0$, NS, respectively. Overall, this *post-hoc* analysis showed a differential improvement in male and female patients on non-motor tests that favored abilities typically performed better by healthy individuals of their sex.

Post-hoc exploratory tests were also conducted to determine whether atypical (risperidone) and typical antipsychotics (haloperidol) have differential effects on

cognitive response in female patients. There were too few patients on the other antipsychotics to examine their effects on cognition. We conducted two multi-factor ANOVAs in a subset of female patients comparing the effects of risperidone to haloperidol over Time (baseline, 5-weeks) on 'female' and then 'male' tests. There were 12 women receiving risperidone treatment ($M = 3.31$ mg/day, $SD = 2.95$) and 13 receiving haloperidol treatment ($M = 3.67$ mg/day, $SD = 3.77$). Parental SES, age, and IQ levels were not significantly different in the two treatment groups. Female patients in this subanalysis, as in the full analysis, showed an overall improvement following treatment on 'female' tests, $F(1, 23) = 6.77$, $p = 0.02$. Notably, although the sample size is small, the pattern of effects was similar for both medication treatments, $F(1, 23) < 1.00$, NS,

and for the interaction between Group and Time on 'female' tests, $F(1, 23) < 1.00$, NS. This smaller group of female patients did not show a significant decrease in performance on 'male' tests following treatment though the effect was in the expected direction, $F(1, 23) = 1.17$, $p = 0.29$, and was similar in both treatment groups, $F(1, 23) = 1.37$, and for the interaction between Group and Time on 'male' tests, $F(1, 23) < 1.00$, NS. Overall this exploratory analysis revealed no differential cognitive response to risperidone and haloperidol treatment in female patients.

Analyses of Clinical Ratings

There was no sex difference in clinical response to treatment, although there was evidence of significant improvement across all patients on each clinical measure: BPRS, $F(1, 54) = 48.55$, $p < 0.001$; SAPS, $F(1, 54) = 81.43$, $p < 0.001$; SANS, $F(1, 54) = 27.95$, $p < 0.001$; and HAM-D, $F(1, 54) = 28.76$, $p < 0.001$, and GAF, $F(1, 54) = 102.40$, $p < 0.001$ (Table 3). Type of antipsychotic (typical, atypical) did not differentially affect patients on any clinical outcome measures ($p > 0.05$). Male and female patients did not differ on their baseline clinical measurements on the BPRS, SAPS, SANS, and the HAM-D, and GAF, $t(56) = -0.09$, NS, $t(56) = -0.44$, NS, $t(56) = 0.62$, NS, $t(56) = 0.04$, NS, and $t(56) = -0.34$, NS, respectively. The improvement of female patients on 'female' tests positively correlated with their clinical improvement on the GAF ($r = 0.44$) and SANS ($r = -0.43$). The decrease in performance of female patients on 'male' tests did not correlate with any measure of clinical improvement ($p > 0.05$). The increase in performance of male patients on 'male' tests, once tests affecting motor control were removed did not correlate significantly with measures of clinical improvement ($p > 0.05$).

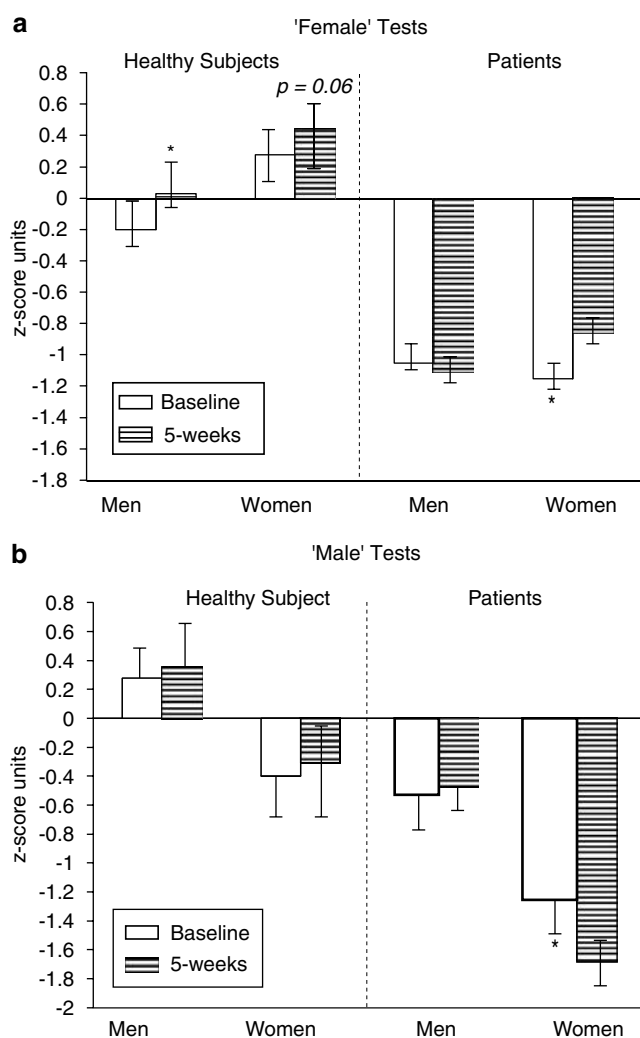


Figure 1 Mean z-scores (+ SE) for healthy subjects and patients as a function of Time and Sex on (a) 'female' tests and (b) 'male' tests. Note. The post hoc analysis showed a significant Group \times Sex \times Time \times Test interaction, $F(1, 105) = 4.75$, $p = 0.04$. Asterisks denote significant pairwise difference in within-subject change over time. Female patients showed significant improvements on 'female' tests and a decrease in 'male' tests over time. Male patients did not show significant improvement on 'male' or 'female' tests. Composite scores were first created by computing within-test z-scores followed by between-test z-scores. Therefore, the composite scores will not sum to 0 as typically expected.

DISCUSSION

In the present study, we sought to extend the investigation of cognitive sex differences in schizophrenia in a novel way by grouping cognitive tests according to the direction of the typical sex difference, an approach used in studies of hormonal effects on behavior in other clinical conditions.

Table 3 Mean and Standard Deviations for the Clinical Outcome Measures for Patients as a Function of Sex and Time

Clinical measures	Male patients		Female patients	
	Baseline	5-weeks	Baseline	5-weeks
BPRS	51.31 (10.17)	40.42 (10.65)	51.55 (9.16)	36.57 (9.18)
SAPS	9.50 (3.11)	5.40 (3.43)	9.91 (3.87)	4.29 (3.26)
SANS	13.86 (3.44)	11.66 (3.53)	13.27 (3.61)	11.00 (3.07)
HAM-D	21.47 (9.70)	13.97 (7.47)	21.36 (11.64)	12.52 (5.65)
GAF	29.75 (6.29)	41.14 (8.21)	30.36 (7.23)	43.43 (7.74)

BPRS, Brief Psychiatric Rating Scale; SAPS, Schedule for assessment of positive symptoms; SANS, the schedule for assessment of negative symptoms; HAM-D, 24-item Hamilton Depression Rating Scale; GAF, Global Assessment of Functioning (higher scores indicate better functioning).

We also sought to explore how performance on these 'male' and 'female' tests changed following antipsychotic treatment. Our hypothesis that women with schizophrenia would show a post-treatment, selective improvement in 'female' tests and a decline on 'male' tests, was supported. Most importantly, we also found support for our prediction that this selective positive response would be greater than any improvement observed in male patients on cognitive tests that favor males or that favor females. These findings of greater sex-specific cognitive enhancement in female compared to male patients indicated effects beyond a similar 'normalization' of sex differences in male and female patients post-treatment.

A series of *post-hoc* analyses addressed other issues relevant to the primary findings of interest. One limitation of the present study is that results were based on an observational study and therefore we had no experimental control over what medications women were receiving. In a *post-hoc* analysis, we detected no differences between risperidone and haloperidol on cognitive performance in female patients although our small sample size provided limited statistical power. Recent studies reported that conventional and atypical prolactin-sparing antipsychotics had similar effects on estrogen levels in women with schizophrenia (Bergemann *et al*, 2005a; Canuso *et al*, 2002) but other studies report differing effects on other endocrine systems, including cortisol and adrenocorticotrophic hormone (Cohrs *et al*, 2004, 2006; Tandon and Halbreich, 2003). A second concern was that extrapyramidal side effects post-treatment might influence the pattern of findings, because motor tests were included in the 'male' and 'female' domains. The pattern of effects was generally the same in analyses excluding motor tests, but the findings are limited by the availability of only one non-motor male test, the JOLO. Future studies with more 'male' tests are needed to re-examine this issue. A third issue was the relationship between sex and cognitive and clinical responses to treatment. There was no sex difference in clinical response to treatment that might account for the sex-specific effects on sexually dimorphic cognitive tests. Clinical improvement post-treatment was significantly related to improvement on 'female' tests among female patients, but unrelated to improvements on the JOLO in male patients and unrelated to declines on 'male' tests among females.

Multiple factors may account for the observed patterns of cognitive changes in female and male patients following treatment. The 'female' tests—the California Verbal Learning Test, Digit Symbol, Grooved Pegboard, and Controlled Oral Word Associations Test—draw on different neural substrates for successful performance, suggesting that the treatment response is not owing to a common underlying neural network but to their common sensitivity to sex-related differences. Previous studies linking these tests together in 'male' and 'female' composite scores have shown changes in performance associated with sex steroid hormones, for example across the menstrual cycle (Hampson, 1990) or with testosterone supplementation (Slabbekoorn *et al*, 1999). The decline in visuospatial abilities in female patients following treatment is notable in light of a previous report in premenopausal women showing that as estradiol levels increased, performance on

visuospatial tests decreased and performance on tests that show a female advantage increased (Maki *et al*, 2002; Hampson, 1990).

One explanation of the present findings is that antipsychotic treatment may have facilitated a normalization of estrogen's activational regulation of neurophysiology and cognition, with consequent positive effects on verbal memory, visual scanning, and psychomotor speed, and negative effects on visuospatial skills in female patients. This explanation cannot be directly tested, because we did not have direct sampling of estrogen in the periphery of the brain, nor a direct measure of estrogen's ability to regulate neurophysiological systems. However, the explanation is parsimonious in light of what is known about the neurobiology of sex differences in cognition, particularly about the differential effects of endogenous estrogen on 'male' and 'female' abilities (Maki *et al*, 2002; Hampson, 1990). The role of estrogen in modulating the clinical course of schizophrenia has been supported in recent clinical trials showing improved clinical response to antipsychotic treatment with exogenous estrogen therapy (Akhondzadeh *et al*, 2003; Bergemann *et al*, 2005b; Kulkarni *et al*, 2001; Louza *et al*, 2004). Additionally, fluctuations in endogenous estrogen across the menstrual cycle are associated with fluctuations in symptoms in women with schizophrenia, with lower severity in the midluteal (high estrogen) phase compared with the follicular low estrogen phase (Hallonquist *et al*, 1993; Thompson *et al*, 2000). Sex differences in other factors, particularly sex-linked genetic factors and other sex hormones, may also contribute to the pattern of findings. Regardless of the mechanism of the observed sex differences in treatment outcome, this pattern of effects in the present study suggests that sex may influence which cognitive abilities improve or decrease with antipsychotic treatment. Such effects may not have been seen in earlier studies when men or women are pooled for analyses, or when tests with different male and female advantages are combined into composite scores according to cognitive domain (eg, executive function).

This study had some notable limitations. Our sample of healthy female subjects included naturally cycling women and women receiving oral contraceptives, whereas women with schizophrenia were not receiving oral contraceptives. Although this would confound a comparison between patients and controls, the primary finding of interest was the different clinical response between male and female patients, and no female patients were receiving oral contraceptives. Secondly, our 'female' and 'male' composites had moderate internal consistency levels, which were somewhat lower for 'male' tests (0.51). This difference in internal consistency might contribute to the finding of a smaller effect size for female patients on the 'male' tests compared with the 'female' tests. Adding additional sexually dimorphic tests to enhance the 'male' composite index may help to increase internal consistency and potentially enhance detection of change on male tests. Lastly, replication of the study is needed.

In summary, the present study examined sex differences in the cognitive response to antipsychotic treatment in a relatively large sample of antipsychotic-naïve first-episode schizophrenia patients. When neuropsychological abilities were parsed into male and female domains according to

the presence and direction of a normative sex difference, a differential pattern of cognitive response to antipsychotic treatment by male and female schizophrenia patients was evident. Following treatment, female patients demonstrated a significant cognitive improvement on 'female' tests and a significant cognitive decline on 'male' tests. Male patients did not change significantly on 'female' tests and showed improvement on 'male' tests when motor tests, which are often adversely affected by treatment via effects on extrapyramidal systems, were excluded from the analysis. A parsimonious explanation for these effects is that during acute psychosis, there may be a reduced regulation of neural physiology by estrogen. Treatment led to a recovery of typical patterns of sex differences in cognitive abilities, perhaps by normalizing the ability of estrogen to have activational effects on neural systems. Although some other factor or set of factors might account for the observed sex-difference in cognitive outcome of antipsychotic treatment, the observation itself is important for understanding cognition-enhancing effects of antipsychotics and it implies a need to take potential sex differences into account in planning future studies comparing approaches for reducing cognitive deficits. Both for this purpose and to learn more about sex differences in schizophrenia and the effects of treatment upon them, cognitive studies may benefit from including more 'female' and 'male' biased tests in designing test batteries for protocols investigating the cognitive effects of drug treatments.

ACKNOWLEDGEMENTS

This publication was supported by funds received from MH62134, MH45156, MH01433, NIH/NCRR/GCRC Grant M01 RR00056, and NARSAD. We thank Drs Cameron Carter and Debra Montrose and the staff of the Clinical Core of the Pittsburgh Center for the Neuroscience of Mental Disorders for providing subject recruitment services and independent diagnostic evaluations and clinical ratings.

REFERENCES

- Akhondzadeh S, Nejatisafa AA, Amini H, Mohammadi MR, Larijani B, Kashani L *et al* (2003). Adjunctive estrogen treatment in women with chronic schizophrenia: a double-blind, randomized, and placebo-controlled trial. *Prog Neuropsychopharmacol Biol Psychiatry* 27: 1007–1012.
- Albus M, Hubmann W, Mohr F, Scherer J, Sobizack N, Franz U *et al* (1997). Are there gender differences in neuropsychological performance in patients with first-episode schizophrenia? *Schizophr Res* 28: 39–50.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Press: Washington, DC.
- Andreasen NC (1994a). *Scale for the Assessment of Positive Symptoms (SAPS)*. University of Iowa Press: Iowa City.
- Andreasen NC (1994b). *Scale for the Assessment of Negative Symptoms (SANS)*. University of Iowa Press: Iowa City.
- Benton AL, Hamsher K, Varney NR, Spreen O (1983). *Contributions to Neuropsychological Assessment: A Clinical Manual*. Oxford University Press: New York.
- Bergemann N, Mundt C, Parzer P, Jannakos I, Nagl I, Salbach B *et al* (2005a). Plasma concentrations of estradiol in women suffering from schizophrenia treated with conventional vs atypical antipsychotics. *Schizophr Res* 73: 357–366.
- Bergemann N, Mundt C, Parzer P, Pakrasi M, Eckstein-Mannspurger U, Haisch S *et al* (2005b). Estrogen as an adjuvant therapy to antipsychotics does not prevent relapse in women suffering from schizophrenia: results of a placebo-controlled double-blind study. *Schizophr Res* 74: 125–134.
- Burton LA, Henninger D, Hafetz J (2005). Gender differences in relations of mental rotation, verbal fluency, and SAT scores to finger length ratios as hormonal indexes. *Dev Neuropsychol* 28: 493–505.
- Canuso CM, Goldstein JM, Wojcik J, Dawson R, Brandman D, Klibanski A *et al* (2002). Antipsychotic medication, prolactin elevation, and ovarian function in women with schizophrenia and schizoaffective disorder. *Psychiatry Res* 111: 11–20.
- Chavez EL, Trautt GM, Brandon A, Steyaert J (1983). Effects of test anxiety and sex of subject on neuropsychological test performance: Finger Tapping, Trail Making, Digit Span and Digit Symbol tests. *Percept Mot Skills* 56: 923–929.
- Cohrs S, Pohlmann K, Guan Z, Jordan W, Meier A, Huether G *et al* (2004). Quetiapine reduces nocturnal urinary cortisol excretion in healthy subjects. *Psychopharmacology (Berl)* 174: 414–420.
- Cohrs S, Roher C, Jordan W, Meier A, Huether G, Wuttke W *et al* (2006). The atypical antipsychotics olanzapine and quetiapine, but not haloperidol, reduce ACTH and cortisol secretion in healthy subjects. *Psychopharmacology (Berl)* 185: 11–18.
- Delis D, Kaplan E, Ober BA (1987). *California Verbal Learning Test*. Adult version manual The Psychological Corporation: San Antonio.
- Endicott J, Spitzer RL, Fleiss JL, Cohen J (1976). The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 33: 766–771.
- Goldberg TE, Gold JM, Torrey EF, Weinberger DR (1995). Lack of sex differences in the neuropsychological performance of patients with schizophrenia. *Am J Psychiatry* 152: 883–888.
- Goldstein JM, Seidman LJ, Goodman JM, Koren D, Lee H, Weintraub S *et al* (1998). Are there sex differences in neuropsychological functions among patients with schizophrenia? *Am J Psychiatry* 155: 1358–1364.
- Grigoriadis S, Seeman MV (2002). The role of estrogen in schizophrenia: implications for schizophrenia practice guidelines for women. *Can J Psychiatry* 47: 437–442.
- Gur RC, Ragland JD, Moberg PJ, Bilker WB, Kohler C, Siegel SJ *et al* (2001). Computerized neurocognitive scanning: II. The profile of schizophrenia. *Neuropsychopharmacology* 25: 777–788.
- Haas G, Sweeney JA, Hien DA, Goldman D, Deck M (1991). Gender differences in schizophrenia (abstract). *Schizophr Res* 4: 277.
- Halari H, Mehrotra R, Sharma T, Ng V, Kumari V (2006). Cognitive impairment but preservation of sexual dimorphism in cognitive abilities in chronic schizophrenia. *Psychiatry Res* 141: 129–139.
- Hallonquist JD, Seeman MV, Lang M, Rector NA (1993). Variation in symptom severity over the menstrual cycle of schizophrenics. *Biol Psychiatry* 33: 207–209.
- Hamilton M (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23: 56–62.
- Hampson E (1990). Variations in sex-related cognitive abilities across the menstrual cycle. *Brain Cogn* 14: 26–43.
- Hill SK, Keshavan MS, Thase ME, Sweeney JA (2004). Neuropsychological dysfunction in antipsychotic-naïve first-episode unipolar psychotic depression. *Am J Psychiatry* 161: 996–1003.
- Hoff AL, Wieneke M, Faustman WO, Horon R, Sakuma M, Blankfeld H *et al* (1998). Sex differences in neuropsychological functioning of first-episode and chronically ill schizophrenic patients. *Am J Psychiatry* 155: 1437–1439.
- Kimura D (1992). Sex differences in the brain. *Sci Am* 267: 118–125.
- Kimura D, Hampson E (1993). Neural and hormonal mechanisms mediating sex differences in cognition. In: Vernon PA (ed).

- Biological Approaches to the Study of Human Intelligence*. Ablex Publishing Corp: New Jersey. pp 375–397.
- Kopala L, Clark C, Hurwitz TA (1989). Sex differences in olfactory function in schizophrenia. *Am J Psychiatry* **146**: 1320–1322.
- Kramer JH, Delis DC, Daniel M (1988). Sex difference in verbal learning. *J Clin Psychol* **44**: 915.
- Kramer JH, Delis DC, Kaplan E, O'Donnell L, Prifitera A (1997). Developmental sex differences in verbal learning. *Neuropsychology* **11**: 577–584.
- Kulkarni J, Riedel A, de Castella AR, Fitzgerald PB, Rolfe TJ, Taffe J et al (2001). Estrogen—a potential treatment for schizophrenia. *Schizophr Res* **48**: 137–144.
- Lewine RR, Walker EF, Shurett R, Caudle J, Haden C (1996). Sex differences in neuropsychological functioning among schizophrenic patients. *Am J Psychiatry* **153**: 1178–1184.
- Louza MR, Marques AP, Elkis H, Bassitt D, Diegoli M, Gattaz WF (2004). Conjugated estrogens as adjuvant therapy in the treatment of acute schizophrenia: a double-blind study. *Schizophr Res* **66**: 97–100.
- Maki PM, Rich JB, Rosenbaum RS (2002). Implicit memory varies across the menstrual cycle: estrogen effects in young women. *Neuropsychologia* **40**: 518–529.
- Mann VA, Sasanuma S, Sakuma N, Masaki S (1990). Sex differences in cognitive abilities: a cross-cultural perspective. *Neuropsychologia* **28**: 1063–1077.
- McCurry SM, Gibbons LE, Uomoto JM, Thompson ML, Graves AB, Edland SD et al (2001). Neuropsychological test performance in a cognitively intact sample of older Japanese American adults. *Arch Clin Neuropsychol* **16**: 447–459.
- Morrison MW, Gregory RJ, Paul JJ (1979). Reliability of the Finger Tapping Test and a note on sex differences. *Percept Mot Skills* **48**: 139–142.
- Overall J, Gorman D (1962). The Brief Psychiatric Rating Scale. *Psychol Rep* **10**: 799–812.
- Perlick D, Mattis S, Stastny P, Teresi J (1992). Gender differences in cognition in schizophrenia. *Schizophr Res* **8**: 69–73.
- Rahman Q, Wilson GD (2003). Large sexual-orientation-related differences in performance on mental rotation and judgment of line orientation tasks. *Neuropsychology* **17**: 25–31.
- Reitan RM, Wolfson D (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. Neuropsychology Press: Tucson.
- Riva D, Benton A (1993). Visuospatial judgment: a cross-national comparison. *Cortex* **29**: 141–143.
- Ruff RM, Parker SB (1993). Gender- and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the Finger Tapping and Grooved Pegboard Tests. *Percept Mot Skills* **76**(Pt 2): 1219–1230.
- Schmidt SL, Oliveira RM, Rocha FR, breu-Villaca Y (2000). Influences of handedness and gender on the grooved pegboard test. *Brain Cogn* **44**: 445–454.
- Seidman LJ, Goldstein JM, Goodman JM, Koren D, Turner WM, Faraone SV et al (1997). Sex differences in olfactory identification and Wisconsin Card Sorting performance in schizophrenia: relationship to attention and verbal ability. *Biol Psychiatry* **42**: 104–115.
- Shimoyama I, Ninchoji T, Uemura K (1990). The finger-tapping test: a quantitative analysis. *Arch Neurol* **47**: 681–684.
- Slabbekoorn D, van Goozen SH, Megens J, Gooren LJ, Cohen-Kettenis PT (1999). Activating effects of cross-sex hormones on cognitive functioning: a study of short-term and long-term hormone effects in transsexuals. *Psychoneuroendocrinology* **24**: 423–447.
- Snow WG, Weinstock J (1990). Sex differences among non-brain-damaged adults on the Wechsler Adult Intelligence Scales: a review of the literature. *J Clin Exp Neuropsychol* **12**: 873–886.
- Spitzer RL, Williams JB, Gibbon M, First MB (1992). The structured clinical interview for DSM-III-R SCID). I: history, rationale, and description. *Arch Gen Psychiatry* **49**: 624–629.
- Tandon R, Halbreich U (2003). The second-generation 'atypical' antipsychotics: similar improved efficacy but different neuroendocrine side effects. *Psychoneuroendocrinology* **28**(Suppl 1): 1–7.
- Thompson K, Sergejew A, Kulkarni J (2000). Estrogen affects cognition in women with psychosis. *Psychiatry Res* **94**: 201–209.
- Thurstone L, Thurstone T (1938). *Primary Mental Abilities*. University of Chicago Press: Chicago.
- Wechsler D (1981). *Wechsler Adult Intelligence Scale-Revised Manual*. The Psychological Corporation: New York.