A Prospective Trial of Customized Adherence Enhancement Plus Long-Acting Injectable Antipsychotic (CAE-L) in Individuals with Schizophrenia or Schizoaffective Disorder at Risk for Treatment Non-Adherence and for Homelessness

Sajatovic M\textsuperscript{1,2}, Levin J\textsuperscript{2}, Ramirez L\textsuperscript{2}, Hahn D\textsuperscript{2}, Tatsuoka C\textsuperscript{1}, Seifi N\textsuperscript{1,2}

\textsuperscript{1}Neurological Outcomes Center, University Hospitals Case Medical Center, Cleveland, OH, \textsuperscript{2}Department of Psychiatry, University Hospitals Case Medical Center, Cleveland, OH

Methodological Question Being Addressed: The effects of CAE-L on adherence, symptoms, and functioning in homeless or recently homeless individuals with schizophrenia/schizoaffective disorder.

Introduction: Treatment non-adherence in people with schizophrenia is associated with relapse and increased rates of homelessness. Building upon the demonstrated utility of long-acting medication, and recent work by this group of investigators in psychosocial interventions to enhance adherence, we conducted a prospective uncontrolled trial of customized adherence enhancement (CAE) plus long-acting injectable antipsychotic (LAI) in 30 homeless or recently homeless individuals with schizophrenia or schizoaffective disorder.

Methods: Participants received monthly CAE combined with LAI (CAE-L) for 6 months. The LAI was haloperidol decanoate. Primary outcomes were adherence and housing status. Secondary outcomes included psychiatric symptoms, functioning, side effects, and hospitalizations.

Results: Mean age of the sample was 41.8 years (SD 8.6) with a high proportion of minorities (90% African-American), relatively high proportion of single/never married individuals (70%) and relatively low levels of education (most did not finish high school). Nearly all (97%) had a history of past or current substance abuse, and nearly all had been to jail or prison (97%). Ten individuals (33%) terminated the study prematurely. Mean end-point dose of LAI was 68.0 mg (SD 21.1). CAE-L was associated with good adherence to maintenance LAI (76% at 6 months) and dramatic improvement in concomitant orally prescribed medication adherence, which changed from missing 46% of prescribed medication at study enrollment to missing only 10% at study end. Mean proportion of time in sub-optimal housing went from 56% in the 6 months prior to study enrollment, to 41 % in the first 3 months of the study and 14% in the last 3 months of the study (p= .001). There were significant improvements in psychiatric symptoms and functional status. The major side effect with LAI was akathisia, which lead to LAI discontinuation in one individual.

Conclusion: While interpretation of study findings must be tempered by the methodological limitations, CAE-L appears to be associated with improved adherence, symptoms, and functioning in homeless or recently homeless individuals with schizophrenia/schizoaffective disorder. Proportion of time in sub-optimal housing was decreased. There is a need for additional research on effective and practical approaches to improving health outcomes for homeless people with schizophrenia.

Support for this study was provided by The Reuter Foundation and CTSC (DCRU) grant# UL1 RR024989.

Disclosures: The authors report no conflicts of interest for this work

Brain-Derived Neurotrophic Factor Genotype Modulates Amygdala Habituation in Borderline Personality Disorder

Perez-Rodriguez M\textsuperscript{1,2,3}, New A\textsuperscript{1,2}, Goldstein K\textsuperscript{1}, Rosell D\textsuperscript{1,2}, Yuan Q\textsuperscript{4}, Zhou Z\textsuperscript{4}, Hodgkinson C\textsuperscript{4}, Goldman D\textsuperscript{4}, Siever L\textsuperscript{1,2}, Hazlett E\textsuperscript{1,2}

\textsuperscript{1}Neurological Outcomes Center, University Hospitals Case Medical Center, Cleveland, OH, \textsuperscript{2}Department of Psychiatry, University Hospitals Case Medical Center, Cleveland, OH, \textsuperscript{3}Department of Neurology, University Hospitals Case Medical Center, Cleveland, OH, \textsuperscript{4}Department of Psychiatry, Case Western Reserve University, Cleveland, OH
Methodological Question Being Addressed: Can the imaging-genetics approach be used to elucidate the genetic underpinnings of the amygdala habituation deficit found in borderline personality disorder (BPD)? Specifically, do Brain-Derived Neurotrophic Factor (BDNF) genotypes modulate amygdala reactivity and habituation to emotional stimuli in BPD?

Introduction: Borderline personality disorder is a psychiatric disorder characterized by emotion-processing abnormalities. Elucidating its underlying neural systems and genetic modulators and fractionating it into neurobiologically defined pathophysiologic subtypes is crucial for refining testable models and developing personalized treatments for emotion dysregulation. Amygdala hyper-reactivity and deficient habituation are putative endophenotypes of abnormal emotion processing in BPD, which are under genetic modulation by brain-derived neurotrophic factor (BDNF) variants. The Met allele of the Val66Met SNP of the BDNF gene increases amygdala reactivity and impairs extinction learning, a phenomenon closely related to habituation. We aimed to use an imaging-genetics framework to examine for the first time in BPD patients the impact of BDNF Val66Met genotypes on amygdala habituation to repeated emotional and neutral pictures. We hypothesized that BDNF 66Met-carrying BPD patients would have a deficit in amygdala habituation to repeated unpleasant emotional pictures compared to non-Met carrying BPD patients and Met-carrying- and Non-Met carrying SPD and HCs.

Methods: We employed event-related functional magnetic resonance imaging (fMRI) in 57 subjects (19 unmedicated BPD and 18 schizotypal personality disorder patients and 20 healthy controls) during a task involving viewing of unpleasant, neutral, and pleasant pictures, presented twice. Amygdala responses were examined with a mixed-model multivariate MANOVA including BDNF Val66Met SNP genotype (Met-carriers vs. Non-Met carriers).

Results: A significant Diagnostic group×Genotype (BDNF Val66Met SNP Met- vs. Non-Met-carriers)×Picture type (unpleasant, neutral, pleasant)×Picture repetition (Novel/Repeat)×Time interaction indicated that Met-carrying BPD patients (but not Met-carrying SPD patients or HCs) showed exaggerated amygdala reactivity to repeated, but not novel, unpleasant pictures, representing a habituation deficit.

Conclusions: Using an imaging-genetics approach, we characterized for the first time the genetic underpinnings of an amygdala habituation deficit to emotional stimuli in BPD, which is restricted to those carrying the BDNF 66Met allele. This finding points to BDNF modulators as a novel therapeutic avenue for BPD, a disorder which lacks FDA-approved medications.

Disclosures: The authors report no conflicts of interest for this work.

Acknowledgments: This research was supported by grant R01MH073911 to EAH, a Merit Award from the Department of Veterans Affairs to ASN (9001-03-0051), grant UL1RR029887 from the National Center for Research Resources, National Institutes of Health, and the Department of Veterans Affairs (VISN3) Mental Illness Research, Education, and Clinical Center (MIRECC). The funding sources had no role in the design of the study, collection and analysis of data and decision to publish. Part of the data presented in this article was presented as a poster at the Society of Biological Psychiatry’s 68th Annual Scientific Convention & Meeting, San Francisco, CA. May 2013 (629 - Impaired Extinction of Amygdala Response to Unpleasant Pictures in Borderline Personality Disorder Associated with BDNF genotype).

3 Prospective SIB Assessment: An Internet Survey of Pharmaceutical Sponsor Practices

ISCTM Suicidal Ideation and Behavior Assessment Workgroup: Chappell P1, Mahabreshwarkar A2 (Co-Chairs), Alphs L3, Bangs M4, Butler A5, Dubrava S1, Feltner D6, Greist J7, Harkavy-Friedman J8, Lenderking W6, Mundt J10, Palumbo D1, Stewart M1

1Pfizer, 2Takeda, 3Janssen Scientific, 4Lilly, 5Bracket, 6Abbvie, 7Healthcare Technology Systems, 8American Foundation for Suicide Prevention, 9Evidera, 10Center for Telepsychology

Methodological Question Being Addressed: What are current approaches of clinical trial sponsors in prospective suicidal ideation and behavior (SIB) assessments and challenges encountered?

Introduction: After the initial US FDA draft guidance was released (August 2010) [1], prospective assessments of SIB were broadly implemented in industry-sponsored clinical studies. Complementing a previous survey of site experiences and attitudes toward prospective assessment of SIB in clinical trials [2], the ISCTM SIB Assessment Workgroup (ISAW)
conducted a survey to obtain systematic information on sponsor practices and experiences in implementing SIB assessments in industry-sponsored clinical trials.

Methods: Potential challenges and issues were identified anecdotally in discussions with stakeholders within and outside the ISAW. Based on these discussions, a 30-item survey, including questions on respondents’ background, was developed through an iterative process of review by the entire group. The survey was sent to 1447 industry employees at 178 pharmaceutical companies from the ISCTM membership mailing list and a contact mailing list maintained by a vendor that provides clinical trial scientific services. Representatives from CROs, vendors, or academic and government institutions were excluded. The survey was implemented using the online survey software SurveyMonkey™.

Results: A total of 132 responses from 50 different companies were collected over five weeks (response rate of 9.1%). Respondents who had no involvement in SIB assessments in clinical trials were excluded from further analyses, leaving a total of 89 evaluable responses, representing 39 companies. Approximately half of the respondents worked at large and a third, at mid-sized companies. About 40% were physicians (23% psychiatrists), and the majority (81%) worked in clinical development.

Common factors used in deciding whether or not to include an SIB assessment in a clinical trial were: psychiatric or neurologic drug product; CNS activity; the disease and patient population under study; and regulatory announcements and policies. The most common indications for which SIB monitoring was performed included (in descending order): schizophrenia, depression, bipolar disorder, Alzheimer’s disease or other dementia, anxiety, ADHD, pain, and Mild Cognitive Impairment (MCI).

Ninety-nine percent of respondents reported SIB assessments were performed at both screening and baseline visits. The most common look-back periods used at the screening visit were 1 month, 6 months, 1 year, and lifetime; and, at baseline, the interval since the screening visit. The majority of respondents (95%) reported utilizing the Columbia-Suicide Severity Rating Scale (C-SSRS); but 20% and 10% also reported using the Sheehan Suicide Tracking Scale and the InterSePT Scale for Suicide Thinking, respectively. The top five challenges in implementing SIB assessments included: cross-cultural differences in acceptance of SIB assessments (40%), obtaining adequate baseline history (36.8%), obtaining translations in relevant languages (35%), investigator/rater discomfort with asking about SIB (32%), and inadequate training of raters to administer SIB ratings (30%).

Conclusions: Among the sponsors surveyed, the implementation rate of SIB assessment in CNS studies is very high. Most have used the C-SSRS. Challenges regarding standardization of retrospective assessment timeframes and differing approaches to summarization and analysis of SIB-related study data were frequently reported. Supplemental identification of suicide-related information by study sponsors remains common. These results suggest that inconsistent reports of SIB within study datasets are likely and integration of data across studies may still be problematic. Limitations of this study include non-random sampling to identify respondents, a relatively low survey response rate and the possibility that multiple respondents may have come from the same sponsor.

References:

Disclosures: One or more authors report potential conflicts which are described in the program.

4 Past and Future Trends of Clinical Trials for Autism Spectrum Disorders: Analysis of the ClinicalTrials.gov Results Database

Levine S¹, Kolevzon A², Reichenberg A²

¹University of Haifa, Israel, ²Department of Psychiatry and Seaver Center for Autism Treatment and Research, Icahn School of Medicine at Mount Sinai NY

Methodological Question Being Addressed: Past and Future Trends of Clinical Trials for Autism Spectrum Disorders: Analysis of the ClinicalTrials.gov Results Database.

Objective: To characterize clinical trials registered in ClinicalTrials.gov to inform the future research and treatment of Autism Spectrum Disorders.
Method: Analysis consisted of text parsing and computing statistics of ClinicalTrials.gov data that were publicly available to May 2013 and classified for Autism Spectrum Disorders. Studies were examined by observational vs. interventional study type on participant enrollment features, funders, procedures, methodology, and outcomes used. Forecast analysis of funding trends was computed.

Results: Since its instigation, ClinicalTrials.gov received 356 studies of autism. There were no sex or age group differences by observational vs. interventional study type. Interventionsal studies were statistically significantly more likely to be funded by industry than another source (OR=2.87 [1.25, 6.55]). Drug and (OR=45.88 [11.06,190.36]) and behavioral procedures (OR=30.49 [4.17, 223.04]) were more prevalent in interventional than observational trials. A placebo was included in 23.1% of interventional trials. Outcomes that were more commonly used in interventions than observational study designs were: scales and batteries (54.21%; OR=47.95, [11.56,198.97]) and cognitive outcomes (34.07%, OR=10.20 [3.62, 28.73]). Although, dropout is a common outcome in other fields it was a primary outcome in one intervention. Forecast analysis of annual trial registration rates projected an increase in trial registration in the next 5 years, which was not attributed to a single funding source.

Conclusions: Consolidation of outcome measures and data integration is necessary.

5 Use of Modeling and Simulations in Selecting ELND005 (Scyllo-Inositol) Dose in a Bipolar Maintenance Study

Liang E, Kurth M, Bairu M, Abushakra S

Elan Pharmaceuticals, LLC, Cambridge, MA

Methodological Question Being Addressed: Dose selection for a Phase 2 study of a potential mood stabilizing agent in Bipolar type I patients (BPD) based on population pharmacokinetic (PK), correlations of PK and biomarker/pharmacodynamic (PD), and modeling and simulations.

Introduction: Myo-inositol plays an important role in neuronal phospho-inositol signaling and is known to be elevated in bipolar depression and mania. Myo-inositol reduction is thought to underlie the mood stabilizing effect of lithium (Moore et al, 1999). ELND005 (scylo-inositol, a myo-inositol stereoisomer) is currently being evaluated as a potential adjunctive maintenance treatment of BPD I patients (Study BPD201, NCT01674010). In an Alzheimer disease clinical trial (AD201; Salloway et al., 2011), the 250mg BID dose showed acceptable long-term safety and lowered myo-inositol levels by ~ 40%, similar to known effects of lithium. Post hoc analysis suggested that this dose of ELND005 seemed to reduce emergence of depression and anxiety over 78 weeks in Mild AD patients (Abushakra et al., 2012), and the reduction of myo-inositol (PD effect) correlated with decreased emergence of depression and anxiety (Tariot et al., AAIC 2012). Plasma ELND005 exposures corresponding to this PD effect in AD patients are considered the target exposures for the BPD study. Based on the population pharmacokinetic (PopPK) model previously developed (Liang et al., AAIC 2011). Modeling and Simulations (M&S) in BPD patients were performed to estimate likely ELND005 exposures in this population. A dose of 500mg BID projected to provide a comparable ELND005 plasma exposure was hence selected for this relatively younger BPD study population (500mg BID). We herein compare actual versus simulation-predicted plasma levels from the ongoing BPD study.

Methods: PopPK analyses and model development of ELND005 were implemented within NONMEM VI or 7.1.0 with Intel® Visual Fortran. Simulations were previously performed to characterize the effects of dosing regimen, inter-individual variability, and uncertainty on expected plasma PK profiles. Simulations based on 30 subjects per dose group, and 50 replicates of each dose group were drawn to capture the expected inter-trial variability of exposures, and compared to actual observed concentrations from the ongoing BPD study. This study includes an open label treatment phase for 16 weeks and followed by randomization into placebo-controlled phase for up to 48 weeks (as adjunct to either lamotrigine or valproate). Actual observed plasma concentrations of ELND005 from the open phase are presented and compared those from study AD201, and further included to update PopPK model.

Results: To date, a total of 149 BPD patients provided 690 plasma samples from the open treatment phase in Study BPD201, which were analyzed and reported. Plasma concentrations in BPD patients (median age: 45 years) receiving 500mg BID were superimposable on those of AD patients (median age: 75 years) receiving 250mg BID, and showed no overlap with those in AD patients receiving 1000mg BID. Visual Predictive Check (VPC) confirmed a good quality of fit of the updated PopPK model that included PK data from current available Study BPD201.

Conclusions: The observed PK results in BPD patients matched the PK levels predicted from simulations. This data set supports the underlying assumptions of the PopPK/PD Model of ELND005 in BPD patients. The concordance of observed interim PK data with simulated exposures supports dose selection for this study and for future BPD studies.
Disclosures: Drs. Liang, Kurth, Bairu, and Abushakra are full time Elan employees, and stockholders in Elan Pharmaceuticals.

6 Are Touch Screen Devices Appropriate for Neurocognitive Testing?
Wei J1, Morgan D2, Boyd A2
1UNC School of Pharmacy, 2CNS Vital Signs LLC

Methodological Question Being Addressed: Are Touch Screen Devices Appropriate for Neurocognitive Testing?

Objective: The primary study objective is to evaluate the sensitivity of touch screen device on test result timing compared to keyboard device in neurocognitive testing.

Method: Healthy volunteers were selected to take the CNS Vital Signs (CNSVS) neurocognitive test battery either on the laptop using web browser interface, or on the tablet using web browser interface.
Participants: 15 participants were selected for Group A, 15 control subjects for Group B were selected from normative database age matched to Group A. Demographics: Group A age range from 29 to 54 with mean of 42.5 years old, Group B age range from 29 and 59 with mean of 42.6 years old.

Results: Reaction Time Comparison: using T-test, the recorded reaction time between tablet and laptop revealed significant differences in the VBM’s initial reaction time, VIM’s initial reaction time, and CPT’s correct hits reaction time with p < 0.0003 for all three test reaction times. Although VBM’s and VIM’s delayed reaction time did not reach significant difference, they were all trending toward significant differences with tablet’s mean reaction time greater than laptop’s mean reaction time.
Correct/Incorrect Response Comparison: With the CPT normal subjects are expected to have near perfect scores. The mean difference for CPT correct response is 6.6 with p = 0.0249. A subset of Group A participants using a tablet had extremely low CPT correct response, even though the proctor visually confirmed these participants tapped the touch screen during the correct stimulus. Most of the VBM and VIM’s correct and incorrect scores did not show significant differences. Again using T test, only VBM’s delayed correct pass score showed significant difference, p=0.031, which could have been due to subject perform in the 2 groups.

Conclusions: There are significant differences in device reaction time when comparing touch screen to keyboard. Tablets and other touch screen devices should be used with caution in computer based neurocognitive tests. Researchers and clinicians should understand that they do not provide reliable reaction time when compared to devices with native keyboard.

7 Performance of Adults with Down Syndrome on Cognitive and Behavioral Tests: Baseline Data from Phase2a Study of the Investigational Agent ELND005
Kesslak J1, Lott F2, Fan F1, Bairu M1, Wesnes K3, Abushakra S1
1Elan Pharmaceuticals, Cambridge MA, 2Department of Neurology, University of California, Irvine, 3Bracket Global, Goring-On-Thames, UK

Methodological Question Being Addressed: Identify sensitive and reliable cognitive and behavioral tests for use as key outcomes in clinical trials of potential cognitive enhancers in young adults with Down Syndrome (DS).

Introduction: Young Adults with DS exhibit various levels of cognitive and behavioral impairment and functional disability. Triplication of the APP gene is implicated in the life-long overproduction of β-amyloid in DS brain, which is thought to play a role in neuronal toxicity and increased dementia risk. DS brain also shows increased myo-inositol levels, likely due to triplication of the myo-inositol active transporter. The myo-inositol increase correlates with cognitive impairment, possibly due to dysregulation of neuronal phospho-inositol signaling. In Alzheimer patients (AD), ELND005 (Scyllo-inositol), an endogenous myo-inositol isomer, has shown both amyloid and myo-inositol lowering effects in CSF and brain respectively (Salloway et al. 2011, Tariot et al. 2012). At a dose that showed acceptable long-term safety ELND005 showed beneficial trends on cognition in mild AD (Salloway et al. 2011), and is now being developed as a potential cognitive enhancer in DS (Study ELND005-DS201, NCT # 01791725). To date, no such drugs have been approved for DS, and no consensus exists on optimal cognitive outcome measures for clinical trials.

Methods: In Study DS201, safety, pharmacokinetics and cognitive/behavioral effects of 2 doses of ELND005 and placebo are evaluated over 4 weeks in approximately 20 DS subjects without dementia. Eligibility criteria include age between 18-45, K-BIT non verbal IQ scores ≥ 40, and absence of recent cognitive decline or major psychiatric disorder. Cognitive outcome measures include the Computerized Dementia Rating system (CDR System; Keith et al. 1998 and
Introduction: Alzheimer’s Disease (AD) results in subtle cognitive and biological changes in the brain long before a formal diagnosis of dementia or even mild cognitive impairment (MCI) can be established. The identification of early cognitive and functional impairment is increasingly critical. Traditional neuropsychological and functional measures are often not well suited to identify the subtle changes in cognition that manifest during the earliest stages of AD. To this end, we have been actively involved in developing state-of-the-art measures that are extremely sensitive to very early markers of AD. In addition, we have developed computer-based, remotely deliverable functional task simulations that tap functional abilities to perform important complex daily tasks such as following physician directives, medication management and managing financial transactions.

Methods: We present three studies employing newly developed cognitive and functional measures. The first study, examined medial temporal lobe atrophy (MTA) using MRI scans of 16 individuals with MCI or mild probable AD (mean MMSE score was 24.38+ 3.8) which was related to a test of semantic interference (LASSI-L). The second study, examined the association between brain amyloid load with Florbetapir and LASSI-L performance among 17 non-demented elders with memory complaints. Finally, we evaluated 23 patients with aMCI and 21 patients with probable AD who received three novel technology-based task simulations using: a) an automatic teller machine (ATM); b) a voice menu telephone banking system; and c) interacting with a voice menu to change services from an electric company.

Results: MTA scores were negatively correlated with LASSI-L List A2 Cued Recall (r=-.53; p<.04), first cued recall of List B (r=-.75; p<.001) and second cued recall of List B (r=-.78; p<.001). Among elders with memory complaints, LASSI-L measures sensitive to release from proactive interference was correlated with overall SUVR scores r=-.56 (p<.05), SUVR scores in the Precuneus r=-.60 (p<.01) and Posterior Cingulate r=-.74; p<.01. On the ATM task, 52.2% of aMCI subjects evidenced deficits compared to 95% of mild AD subjects (X2= 7.74; p<.01). On the Electric Company task 11.1% of aMCI subjects exhibited deficits compared to 50% of mild AD X2=3.76; p<.05). On the Banking Task

Conclusions: Performance of the CDR System in DS adults seems to support its utility as key cognitive measure in DS trials. Performance of CDR System and the other behavioral measures in this study, and their potential sensitivity to drug effect will inform the choice of appropriate outcome measures in future clinical trials in adults with DS. Enrollment is this study continues and results will provide information for future studies.

Disclosures: Drs. Kesslak, Fan, Bairu, and Abushakra are full time Elan employees, and stockholders in Elan Pharmaceuticals; Dr. Lott is a consultant and has received research support from Elan for this study; Dr. Wesnes is an employee of Bracket Global and is a consultant for Elan Pharmaceuticals for this Study.

8 The Utility of a Novel Cognitive Paradigm for Identifying and Tracking Subjects at High Risk for Early AD: Implications for Clinical Trials

Curiel R, Loewenstein D, Czaja S, Harvey P, Crocco E

University of Miami Miller School of Medicine, Center on Aging and Department of Psychiatry and Behavioral Sciences, Miami, FL

Methodological Question Being Addressed: Implementation of sensitive cognitive and functional outcome measures is critical to demonstrate the clinical benefit of a drug, particularly in preclinical disease states. The field has a significant need for psychometric tools that enhance clinical trial design (i.e. - inclusion and exclusion criteria), improve measurement of meaningful change, reduce variability associated with testing and informant bias, and aligns with biomarker data used to tailor treatments.

Introduction: The preliminary profile resembles AD more closely than vascular or alpha-synuclein related dementias. CDR System showed good general utility, test-retest correlations on speeded attentional scores (0.93), and on accuracy scores up to 0.74.

Edgar et al. 2008), Rapid Assessment of Developmental Disorders (RADD; Walsh et al. 2007), and WAIS-block design. Behavioral measures include the 12-item Neuropsychiatric Inventory (NPI), and the Vineland Adaptive Behavior Scale, II (VABS-II). To date, 11 patients have provided baseline cognitive and behavioral data.

Results: In the first 11 enrolled patients (8 males, 3 females; 9 Caucasian, 2 African American), mean age was 29.5 years (range 18-40), and mean K-BIT score was 59.3 (40-79). At baseline, RADD mean score was 60.9 (43-70); WAIS block design was 16.4 (6-37); VABS-II mean scores were (communication 172, daily living skills 173.6, socialization 176.1, motor skills 144.9), and NPI total score was 1.9 (0-7). The most common psychiatric symptoms were agitation/aggression in 4/11, aberrant motor behavior in 3/11, and disinhibition in 2/11 subjects. Scores from the CDR System will be presented and contrasted to age-matched controls and to profiles of various types of dementia. DS patients showed large impairments on all aspects of cognitive function (attention, information processing, working memory and episodic non-verbal memory). The preliminary profile resembles AD more closely than vascular or alpha-synuclein related dementias. CDR System showed good general utility, test-retest correlations on speeded attentional scores (0.93), and on accuracy scores up to 0.74.

Disclos
63.6% of aMCI subjects evidenced deficits compared to 100% of mild AD patients (X2=6.40; p<.01).

Conclusions: These results suggest that sensitive measures of vulnerability to proactive interference are strongly related to amyloid load in community-dwelling elders and with medial temporal lobe atrophy in MCI and mild AD subjects on structural MRI. Despite a diagnosis of no dementia, a majority of aMCI subjects evidenced deficits on objective functional measures. The advantages of these novel instruments over traditional paradigms are presented.

9 Linguistic and Cultural Adaptation of a Translated Neurocognitive Battery in Russia, Switzerland and Italy: Preparation for a Program to Delay the Onset of MCI Associated with Alzheimer’s Disease

Atkins A1, Vaughan A1, Turcotte N1, Makeeva O2, Monsch A3, Frisoni G4, Parapini M5, Melikyan Z6, Brewster S7, Oneil J8, Schuster J9, Lymboura M10, Romero H10, Hayden K10, Plassman B10, Welsh–Bohmer K10, Keefe R1,9

1NeuroCog Trials Inc., Durham, NC, 2Neuropsychology Testing Center, Nebbio LLC, Tomsk, Russian Federation, 3University Center for Medicine of Aging Basel, Switzerland, 4IRCCS Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy, 5Zinfandel Pharmaceuticals Inc., Chapel Hill, NC, 6Takeda Development Center, Americas, Deerfield, IL, 7Takeda Development Centre, Europe, London, United Kingdom, 8Joseph and Kathleen Bryan ADRC, Duke University Medical Center, Durham, NC, 9Department of Psychiatry, Duke University Medical Center, Durham, NC

Methodological Question Being Addressed: Cross-cultural adaptation of neurocognitive assessments.

Introduction: Cultural and linguistic accuracy of translated neurocognitive assessments are necessary for appropriate interpretation of outcome data in clinical trials. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) provides guidelines for adapting for patient reported outcomes (Wild et al., 2005, 2009). Building on this guidance, we present a method for cultural adaptation of a neurocognitive battery that incorporates (1) professional translation, (2) in-country pilot testing and expert review processes, and (3) final review and revision of instruments.

Methods: Neurocognitive Battery: Measures included the MMSE, BVMT-R, CVLT-II, Semantic and Lexical Fluency, Digit Span, the Multilingual Naming Test (MiNT), TMT Parts A and B, and the Clock Drawing Test.

Translation: After obtaining permission from copyright holders, each measure underwent formal translation by a professional vendor.

Pilot Testing: Pilot testing in Russia, Switzerland and Italy was conducted to collect feedback from in-country raters (testers) and representative subjects in the populations of interest. Each country sample included 10 subjects (Russian: 5 men, 5 women; German: 5 men, 5 women; Italian: 3 men, 7 women). Ages ranged from 65 to 86.

Expert Review: Following pilot testing, each assessment received formal review by 3 independent neuropsychologists/psychologists in each country. Reviewers provided feedback regarding: construct validity; appropriateness for the target population; recommended revisions.

Final Review: Results were examined by content experts in neurocognitive assessment in the US. Recommendations for changes were made, accommodating cultural adaptations while maintaining consistency of test administration.

Results: Revisions were based on thorough review of feedback received from in-country raters and expert reviewers. Recommended changes included both minor changes to wording or phrasing and more extensive changes to task instructions. In several instances, word substitutions were made to account for differences in word frequency and item familiarity across cultures. In addition, adaptation of stimuli for the MiNT naming task, as well as the list of acceptable responses for several items, was required in all languages. Other changes included minor revisions to the CVLT-II in German, as well as revisions to phrasing and instructions for several instruments in Russian and Italian. All significant changes to each assessment received forward and back translation, as appropriate, prior to incorporation. All changes were then approved by copyright holders prior to finalization.

Conclusions: Cross-cultural adaptation is critical to the success of international clinical trials utilizing neurocognitive endpoints. In-country pilot testing and expert review can provide valuable feedback and serve as the final stage of a formal translation process designed to yield neurocognitive instruments that are linguistically and culturally appropriate. Due to cultural differences in word frequency and item familiarity, this process can be particularly helpful for assessments that rely on word lists originally constructed in a single language. For the present neurocognitive battery, cultural differences in performance are being assessed in an ongoing study to determine instrument reliability, validity, and normative values.

Disclosures: Funding was provided by Takeda Pharmaceutical Company Limited, Zinfandel Pharmaceuticals and
A Comparison of Cross-Cultural Norms for the MATRICS Consensus Cognitive Battery (MCCB)

Stroescu I, Davis V, Atkins A, Keefe R

1NeuroCog Trials, Inc., Durham, NC; 2Department of Psychiatry, Duke University Medical Center, Durham, NC

Methodological Question Being Addressed: Equivalence/differences between norms for culturally adapted/translated versions of the MCCB and the widely used United States norms.

Introduction: The rise in international/global clinical trials that include cognitive, behavioral, and functional outcomes has increased interest in the psychometric characteristics of various measures in different languages and cultures. Multiple important aspects of test theory and construction have to be considered within this context, including fairness and bias, linguistic factors, content and psychometric equivalence, and diagnostic validity (Pedraza & Mungas, 2008). Furthermore, there are genetic, regional, and sociocultural factors that can significantly impact reliability and validity. All of these factors are critical to signal detection in pharmaceutical clinical trials and to the inferences that are drawn from the data. The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) initiative has facilitated the development and application of guidelines for pharmaceutical trials involving neurocognitive outcomes in schizophrenia, including the MATRICS Consensus Cognitive Battery (MCCB), which is now widely used in national and international trials. Several of the individual tests from the MCCB have been used with various international populations, however a comparison of the full battery across languages and cultures will determine the equivalence of the MCCB across languages and the relationship of local normative corrections to signal detection (sensitivity/specificity). We present MCCB data comparing scores that were generated from local (country and language specific) normative adjustments to scores generated from United States (US) norms.

Methods: Neurocognitive Battery: MATRICS Consensus Cognitive Battery (MCCB). Baseline and/or screening MCCB data (composite T-scores and distributions) with normative adjustments from local, language-specific cohorts in the different countries (including India, Russia, China, and Central and South America) were examined and compared to US normative adjustments. The difference in change score distributions was assessed when possible based on data availability.

Results: Preliminary results on 48 patients with schizophrenia using the Hindi-based normative data collected in India show a greater than two standard deviation increase in MCCB composite T-scores relative to the application of US norms. However, the mean (+/- SD) changes in composite T-scores from baseline to week 24 were 7.8 (+/- 7.7) using the Hindi-based norms and 9.3 (+/- 8.9) using the US norms. Additional data will be shown from various countries using MCCB scores analyzed with local and US-based norms to transform scores.

Conclusions: Cross-cultural validity of measures and the availability of appropriate normative data are recommended in international clinical trials with neurocognitive and behavioral endpoints. For the MCCB, the relationship of cultural differences in performance and related psychometric variables/factors regarding signal detection and treatment effects remains a strong consideration for the success of international clinical trials. Preliminary results from India suggest selection of normative data sets may have a considerable impact on the baseline characterization and selection of subjects, although the impact on change scores is less certain. Additional analyses on large data sets are necessary to determine if local norms may affect certain demographic groups within each language and culture, and the extent to which the use of local norms may facilitate treatment signal detection.

Disclosures: Funding for this study was provided by NeuroCog Trials, Inc. The authors are employees of NeuroCog Trials, Inc. Richard Keefe is the founder and CEO of NeuroCog Trials, Inc.

AMPA-Receptor Modulators for Cognitive Impairment in Schizophrenia: Why Are They so Difficult to Develop?

Geerts H, Roberts P, Spiros A

1In Silico Biosciences, Berwyn, PA; 2Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; 3OHSU, Portland, OR

Methodological Question Being Addressed: Understanding the difficulties of developing AMPA-R modulators for cognitive impairment in schizophrenia in order to improve clinical trial design.

Introduction: Positive modulation of AMPA glutamate ion channels has shown promising effects in preclinical animal
models of cognitive enhancement; however clinical trials in cognitive impairment in schizophrenia were not very successful. Also AMPA-R antagonists counter-intuitively have been proposed for cognitive enhancement. We used a humanized quantitative systems pharmacology computer model for studying in detail the impact of AMPA-R modulators on the emergent property of a cortical network that has been calibrated with clinical data on cognition.

**Methods:** The biophysically realistic cortical neuronal network consisted of 80 4-compartment pyramidal cells and 40 two-compartment GABA inhibitory cells with about 10000 synapses, was calibrated using primate electrophysiology data; the effect of GPCR systems such as dopamine, norepinephrine, serotonin and acetylcholine on ion channel conductances was derived from preclinical data and further calibrated using clinical data on working memory. Schizophrenia pathology was introduced as dopamine hypo-activity, NMDA and GABA dysfunction and increased noise; ketamine was implemented as a NR2C specific NMDA-R inhibitor that also increased cortical DA. The output was proportional to the strength of a memory trace and shown to be well correlated with a number of clinical cognitive readouts.

**Results:** We investigated three properties of AMPAR modulators: changing opening time, peak current and closing time and combinations thereof. In all cases the dose-response was non-monotonic and different between the effect on ketamine-induced deficit and in schizophrenia conditions. In all cases with ketamine-induced deficit, the AMPAkines showed an inverse U-shape dose-response with a small dose window of benefit. In most cases with schizophrenia pathology, the dose-response showed first a worsening of the outcome before plateauing at a level not different from placebo. Only in the case where the increase in opening time was less than the increase in closing time did AMPAkines show an improvement but only in a very limited dose-window. Interestingly, conversely, AMPA-antagonists improved cognitive outcome for the inverse combination where opening time was more affected than closing time but always in an inverse U-shape dose-response with a limited effective dose-window.

**Discussion:** This simulation illustrates the complex dose-relationship of AMPA-R modulators on emergent properties of a network that is proportional to a clinical readout of cognitive performance. This is most likely due to complex and non-linear timing relationships between AMPA-R and modulation of NMDA-channels and their effect on the excitation-inhibition balance, leading to complex non-monotonic dose-responses. There is only a limited window of cognitive benefit for specific combinations of changes on opening and closing time of the channel. The simulations also suggested that clinical trials with ketamine are not always predictive for the schizophrenia case. The results highlight the opportunity to use quantitative systems pharmacology modeling as a way to identify possible successful AMPA-R modulators, based on their electrophysiological properties.

**Disclosures:** The authors are employees of In Silico Biosciences, a company that provides knowledge management support for CNS clinical trials.

### 12 Effects of Rater Inconsistency on MCCB Reliability

Rissling A\(^1\), Davis V\(^1\), Walker T\(^{1,2}\), Atkins A\(^1\), Keefe, R\(^{1,2}\)

\(^1\)NeuroCog Trials, INC.; \(^2\)Duke University Medical Center

**The Methodological Question Being Addressed:** Is the performance of adults diagnosed with schizophrenia on the MATRICS Consensus Cognitive Battery (MCCB) sensitive to the effects of rater consistency?

**Introduction:** The MCCB is designed to measure cognitive performance across 7 domains. The MCCB has demonstrated sensitivity to cognitive deficits in schizophrenia, high test-retest reliability and low practice effects. It is often assumed that rater inconsistency decreases the MCCB’s test-retest reliability, potentially reducing sensitivity to treatment effects. Based on this assumption, an effort often is made to maintain rater consistency across assessments in clinical trials. However, the hypothesized benefit of rater consistency on the MCCB has yet to be empirically tested. In order to address this issue, we examined and compared rates of test-retest reliability for MCCB assessments administered by consistent vs. inconsistent raters. If rater consistency is associated with increased reliability, the result would support the necessity of maintaining the same rater at each assessment in pharmaceutical trials.

**Methods:** In order to examine the effects of rater consistency on MCCB reliability, MCCB composite scores were evaluated for 417 schizophrenia patients (mean age=42, range 20-55) at two assessments (T1 and T2) prior to initiation of study drug for two Abbvie clinical trials. Test-retest reliability was evaluated with intraclass correlations that were compared for cases in which the rater was Consistent (n=359) versus Inconsistent (n=58) across visits.

**Results:** The T1 MCCB composite scores for the Consistent (mean=28.19) and Inconsistent (mean=24.48) rater groups were not significantly different (p>.05). High test-retest reliability was exhibited for both the Consistent (ICC = 0.87) and
Inconsistent (ICC= 0.89) rater groups.

**Conclusions:** The results suggest that MCCB test-retest reliability was not sensitive to the effects of rater consistency in this sample. One interpretation of this finding is that requiring rater consistency across assessments may represent an unnecessary constraint in pharmaceutical trials. However, given the limitations of the current analysis, which included a naturalistic design and a relatively low rate of rater variability, more thorough investigation is needed to examine potential interactions between rater consistency and additional factors affecting test-retest reliability of the MCCB, including rater performance and time of testing.

**Disclosures:** Funding for this study was provided by Abbvie, and NeuroCog Trials, Inc. A Rissling is a full-time employee of NeuroCog Trials, Inc., Durham, NC. V Davis is a part-time employee of NeuroCog Trials, Inc., Durham, NC. T Walker is a part-time employee of NeuroCog Trials, Inc., Durham, NC. A Atkins is a full-time employee of NeuroCog Trials, Inc., Durham, NC. R Keefe currently or in the past 3 years has received investigator-initiated research funding support from the Department of Veteran’s Affairs, Feinstein Institute for Medical Research, GlaxoSmithKline, National Institute of Mental Health, Novartis, Psychogenics, Research Foundation for Mental Hygiene, Inc., and the Singapore National Medical Research Council. He currently or in the past 3 years has received honoraria, served as a consultant, or advisory board member for Abbvie, Akebia, Amgen, Astellas, Asubio, AviNeuroChemRar, BiolineRx, Biomarin, Boehringer-Ingelheim, Bristol-Myers Squibb, Eli Lilly, EnVivo, Helicon, Lundbeck, Merck, Mitsubishi, Otsuka, Pfizer, Roche, Shire, Sunovion, Takeda, Targacept. Dr. Keefe receives royalties from the BACS testing battery and the MATRICS Battery (BACS Symbol Coding). He is also a shareholder in NeuroCog Trials, Inc.

**13 Neurocognitive Endophenotypes in Adult Attention-Deficit Hyperactivity Disorder (ADHD)**

Pironti V, Sahakian B

Department of Psychiatry, University of Cambridge, Adult ADHD Clinic Cambridgeshire & Peterborough National Health Service Foundation Trust

**Methodological Question Being Addressed:** Clinical heterogeneity in ADHD, as in the majority of psychiatric disorders, is the norm. This heterogeneity makes it difficult to identify homogenous clinical groups using current nosological instruments (e.g. DSM or ICD). Sample inhomogeneity might clearly lead to spurious drug trial failure (false negative, type II errors). One solution explored by the present study involved detecting neurocognitive endophenotypes in ADHD to determine whether this reduces both group heterogeneity and within sample variability, and therefore potentially reducing false negatives in therapeutic trials.

**Introduction:** Attention-Deficit/Hyperactivity Disorder (ADHD) is a highly heritable neurodevelopmental disorder with childhood onset; characterized by a complex multifactorial pattern of inheritance. Succeeding in unravelling genetic liability for such complex clinical conditions has been challenging as genetic studies have been hard to replicate. Vulnerability markers, also called endophenotypes, are thought to modulate the effect of the genes and the overt behavioural phenotype. Their identification therefore could boost statistical power in molecular genetic studies aiming to identify susceptibility genes linked to the disorder. Other benefits include clarifying the pathophysiology of the disorder and contributing to the development of novel pharmacological treatments tailored on neurocognitive biomarkers. The aim of this study was to identify cognitive and neuroanatomical endophenotypes in adult ADHD.

**Methods:** The study population was twenty adults diagnosed with ADHD according to DSM-IV-TR, twenty unaffected first degree relatives and twenty healthy controls; the three groups being age-matched. All participants underwent a structural magnetic resonance imaging (MRI) scan on a 3T Siemens scanner, and outside the scanner performed two computerized cognitive tests from the Cambridge Neuropsychological Test Automated Battery (CANTAB) assessing motor inhibition (Stop Signal Task – SST) and sustained attention (Rapid Visual Information Processing - RVP). MRI data were processed using voxel-based morphometry with DARTEL algorithm implemented in SPM8; local grey matter and white matter (GM/WM) volumes were used as dependent variables, group as fixed factor, total intracranial volume and age were used as covariates. SPSS version 21 was used for behavioural analyses.

**Results:** Compared to controls Adults with ADHD and their unaffected first degree relatives showed a significant decrease in grey matter volume in the right inferior frontal gyrus (r-IFG) and increased white matter volume in the posterior part of the right inferior fronto-occipital fasciculus (r-IFOF). The ADHD group also shared significant sustained attention impairments with their unaffected first degree relatives compared to the controls. In contrast, response inhibition abilities were not found to be different across groups. Finally, correlation analysis showed dissociations between cognitive functions and neuroanatomy such that there was a positive correlation between sustained attention scores and r-IFG ($r = 0.267; p = 0.039$) volume but not with r- ILOF volume ($r = 0.002; p > 0.985$); while SST scores correlated positively with r- IFOF ($r = 0.294; p = 0.023$) volume but not with r-IFG volume ($r = -0.103; p = 0.435$).
**Conclusions:** In the current study neurocognitive endophenotypes in adult ADHD were identified. Neuroanatomical abnormalities in the right inferior frontal gyrus and in the right inferior fronto occipital fasciculus are familial and might constitute biological vulnerability markers for ADHD. Sustained attention, which is one of the core symptoms of adult ADHD, is also familial and might represent a cognitive vulnerability marker for ADHD. These findings correspond with the recent development of the Research Domain Criteria Approach (RDoC). Using neuroanatomical based endophenotypes found in this study as a neuroscience-based classification in recruiting participants, will improve within subject variability, reduce false positive and improve outcome validity, and will facilitate pharmacological treatment discovery tuned by a biomarker approach.

**Disclosures:** This work was supported by a Medical Research Council (UK) post-graduate studentship to V Pironti, a Wellcome Trust grant to B Sahakian, TW Robbins et al., and completed within the Behavioural and Clinical Neuroscience Institute. BJS consults for Cambridge Cognition, Servier, and Lundbeck; she holds a grant from Janssen/I&J. The remaining author reported no biomedical financial interests or potential conflicts of interest.

14 **Components of Executive Dysfunction in Chronic Schizophrenia and their Relationship to Functional Activities of Daily Living**

Jaeger J1,2, Fichera J3, Maruff P1,4

1CogState, Inc. New Haven, CT, 2Albert Einstein College of Medicine, Bronx, NY, 3Royal Melbourne Institute of Technology, Melbourne Australia, 4Centre for Neuroscience; University of Melbourne, Australia

**Methodological Question Being Addressed:** On a hidden maze learning task paradigm, when errors are classified into those reflecting executive and working memory impairment versus spatial learning impairments, and do rates of these types of errors differentially affect people with schizophrenia and does their frequency correlate with functional activities of daily living?

**Background:** The Groton Maze Learning Task (GMLT) is unique among tests of executive function in that it has been shown to be sensitive to the effects of putative cognitive enhancers in schizophrenia (e.g. Lieberman et al, 2012), is repeatable and is resistant to practice effects. Unlike other tests of executive function which measure the acquisition of the rules of a task, the GMLT measures the application of a set of task rules that are conveyed and practiced prior to testing. The present study investigated distinct types of GMLT performance errors by comparing their frequency in a sample of individuals having schizophrenia to that observed in healthy volunteers. The relationship between these different components of executive dysfunction and functional activities of daily living (FADL) was also examined.

**Methods:** The GMLT was administered to 64 people with schizophrenia and 38 healthy volunteers. All subjects were given the UPSA-2, NART and Brief Psychiatric Rating Scale. GMLT Errors were classified and then divided into ‘rule breaks,’ which represent the ability of working memory to maintain and apply the task rules, and ‘exploratory errors’ reflecting spatial memory impairment. Examples of rule break errors include moving backward or diagonally, double taps, jumps, perseverance, and ‘within-search’ errors in which a location is probed within the search rule but after it has been shown to be incorrect. In contrast, exploratory errors are logical errors that the subject makes while searching for the hidden maze.

**Results:** As expected, total errors were more frequent in the sample with schizophrenia. However a greater proportion of these errors were those rule break errors, reflecting impaired working (rather than spatial) memory. Working (and not spatial) memory impairment were correlated with FADL’s as measured by the UCSD Performance Based Skills Assessment (UPSA-II).

**Conclusions:** The GMLT, a hidden pathway learning task, generates errors that can be classified into executive (working memory) and spatial learning errors. Executive errors are disproportionately observed in schizophrenia and their number is correlated with performance on the UPSA-II performance whereas this functional measure was not correlated with spatial learning errors, supporting the use of this measure in trials testing treatments for schizophrenia cognitive and functional impairments.

15 **Computerized Neuropsychological Tests of Memory: Practice Effects Diminish Sensitivity to Donepezil in Alzheimer’s Disease**

Jaeger J1,2,5,6, Soaita A1, Hannesdottir K1, Maruff P1,4

1CogState, Inc. New Haven, CT, 2Albert Einstein College of Medicine, NY, 3AstraZeneca, 4University of Melbourne, Melbourne, Australia, 5University of Manchester, UK, 6City University of New York, NY

Introduction: Previously we have demonstrated in patients with Alzheimer’s Disease (AD) that sampling cognitive performance on a daily basis for several days provides for more accurate estimations of cognitive functioning both at baseline and after administration of an active treatment (donepezil), and improves the power to detect a drug effect (Jaeger, 2013). This observation follows from demonstrated high levels of day to day cognitive fluctuation in AD (Zettergren, 2011, Jaeger, 2011). The strategy of using daily measurement to capture day to day variability is only feasible if a test is brief. However our work in schizophrenia suggests that the presence of practice effects will also work to obscure the ability to detect the benefit of a treatment (Pietrzak, 2010).

In a study employing multiple daily cognitive measurements at baseline (10x in2wks) and again at three monthly follow ups (5x in1wk) using CogState computerized tests, the CogState One Card Learning (OCL) test detected the effect of donepezil at 2 and 3 months whereas the CPAL task did not. For those analyses, linear mixed models were employed using the mean of all 10 baseline CogState scores as the covariate, the treatment group as the fixed effect, and the site and site by treatment group as random effects. Given the different results of the two computerized memory tests, we sought to determine whether these two tests differed with respect to the effect of repeated administration on test performance as a possible explanation of their differential sensitivity to drug effect.

Methods: Data analyses were conducted from this 3-month RCT comparing donepezil (5 mg for two weeks, followed by 10 mg for the remaining period) to placebo using ADAS-Cog, NTB and multiple daily CogState measurements (as above). The purpose of this investigation was to examine whether practice effects influenced sensitivity to drug effect. A test for trend over time was performed for each CogState task administered at baseline. This was done by testing the contrast \((-5 -4 -3 -2 -1 1 2 3 4 5)\) in the baseline measurement sequence in the mixed model using a t-test. Low p-values indicate a trend over time.

Results: Results indicated that no statistically significant learning effect was detected over 10 repeated measurements on the Cogstate OCL task, whereas the CPAL did show significant learning effects.

Conclusions: One Card Learning, a brief pattern separation learning task, appears not to show learning effects in patients having AD even after 10 repeated daily measurements within a two week testing window. In contrast, the CPAL paradigm was associated with learning effects at least through the seventh administration. OCL and not CPAL showed the benefit of donepezil treatment after two and three months. We hypothesize that the lack of sensitivity of the CPAL paradigm to donepezil may be a consequence of its vulnerability to learning effects.


Wesnes K1,2,3, Annas P4

1 Bracket, Reading, UK, 2 Centre for Human Psychopharmacology, Swinburne University, Melbourne, Australia, 3 Department of Psychology, Northumbria University, Northumbria, UK, 4 Bracket, Wayne, PA

Methodological Question Being Addressed: Which cognitive testing methodologies are most sensitive to the profile of cognitive impairment in Parkinson’s disease?

Background: Dementia develops rapidly in PD, the 8 year prevalence being 78.2% (1). In a recent study 27% of PD patients with Mild Cognitive Impairments (PD-MCI) converted to dementia within 3 years (2). This is notably higher than the 4-6% incidence of Alzheimer’s disease (AD) up to 75 years, as well as the 5-10% annual conversion rate of MCI to AD (3). While the reported incidences of MCI in the general population vary from 5 to 30%, the Mayo Clinic Study using published criteria found it to be 16% (4). The incidence of MCI in 8 cohorts of PD patients was 25.8% (5) and the Movement Disorders Society (MDS) task force review of 48 PD studies identified the incidence to be 26.7% (6). Compared with AD, the more rapid development of dementia in PD suggests that these identified incidences of PD-MCI may be conservative, and the present study sought to determine whether automated tests of cognitive function which combine measures of speed and accuracy identify a higher rate of PD-MCI

Methods: The CDR System has been used extensively in PD. The System is an integrated set of automated tests of attention, information processing, working memory/executive control and episodic memory; both speed and accuracy
scores are captured. CDR System data from 484 PD patients from 6 cohorts were compared to age and gender matched healthy controls. The 2012 MDS Task Force Diagnostic Criteria for PD-MCI Level II (comprehensive assessment) were applied, taking 1 SD as the cut-off. As the guidelines define MCI as 1 to 2 SD below age-matched controls, we also explored higher cut-offs of 1.5 & 2 SD and these will also be reported.

**Results:** Using a cut-off of 1 SD, PD-MCI for attention/information processing was the largest single category at 72.7%, with for example working memory/executive control at 37.4% and episodic memory at 28.8%. 76.4% of the patients qualified for MCI on any one of these three domains.

**Conclusions:** Computerised testing which uses assessments of both accuracy and speed identifies higher rate of MCI than traditional non-automated neuropsychological tests. In PD, speed of information processing, decision making and speed of retrieval of information held in memory are particular impaired, and this may account for the higher rates of MCI detected in this study.

**References:**

**Disclosures:** The author is an employee and stockholder in Bracket, which provides services to the clinical trial industry. One of the services offered is the CDR System, a set of computerised cognitive tests, which was used to gather some of the data presented in this poster.

17 Social Cognition and its Correlates to Functional Outcomes in Schizophrenia

Khan A1,2,3, Lindenmayer J-P2,3,4, Gao L1, Opler M1,4, Lucić L1,5, Rothman B1

1ProPhase LLC, New York, NY, 2Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, 3Manhattan Psychiatric Center, Wards Island, NY, 4New York University School of Medicine, New York, NY, 5Pratt Institute, New York, NY

**Methodological Question Being Addressed:** Do social cognition, neurocognition, and clinical symptoms have a relationship on functional outcomes in patients with schizophrenia?

**Introduction:** Deficits in social functioning, including communication, work, social skills, and community functioning, are a defining feature of schizophrenia. Functional outcomes of schizophrenia are affected by several factors such as social cognition, neurocognition, psychopathology, and clinical outcomes. The multifaceted association among these factors and functional outcome continues to be unclear. Given the significant role of functional outcomes in schizophrenia, there has been increasing importance in factors that may underlie these outcomes. If the characteristics of these factors can be defined, interventions may be developed to improve them, which, in turn, will have a parallel impact on long term functioning and outcome.


**Results:** The overall model fit was χ²=39.6, P<0.14. Fit indexes: Cmin/df=1.25, NFI=0.96, Tucker–Lewis index (Bentler and Bonnet nonformed fit index)=0.97, RMSEA=0.044. Regression weights of the latent variable “Marder Negative Factor” were significant and high (β=0.91) and a substantial amount of variance could be explained by negative symptoms, indicating that the negative factor is a reliable measure of the latent variable. In addition, the regression
weights of the latent variable social cognition to the 3 indicators were moderate and significant (Emotion Recognition DSCB: $\beta=0.77$, ER-40: $\beta=0.46$, and Nonverbal Emotion Identification, $\beta=0.44$). Like social cognition, the latent variable functional assessment explained a substantial amount of variance in the latent variables of working memory ($45\%$).

Impact of social cognition on negative symptoms ($\beta=0.91$) was greater than the direct impact of social functioning ($\beta=0.78$) and functional assessments ($\beta=0.63$).

Conclusions: This study suggests that 49% of negative symptoms could be explained by impaired social cognition and that 49% of social functioning skills could be explained by social cognition. Our findings suggest that social cognition may be an essential target to improve functional outcomes. These findings provide evidence that may help develop novel interventions.

Disclosures: The authors do not have any disclosures to report.

18 Mapping Psychopathology Symptoms in Patients with Schizophrenia to Social Functioning
Khan A$^{1,2,3}$, Harvey P$^4$, Gao L$^1$, Opler M$^{1,5}$

$^1$ProPhase LLC, New York, NY, $^2$Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, $^3$Manhattan Psychiatric Center, Wards Island, NY, $^4$University of Miami Health System, Miami, FL, $^5$New York University School of Medicine, New York, NY

Methodological Question Being Addressed: Would mapping a psychopathology symptom to the construct of social functioning help researchers target specific symptoms and help clinicians individualize treatment interventions when a social functioning measure is not readily available?

Introduction: Patients with schizophrenia usually have psychotic symptoms that significantly impact their social functioning, particularly interpersonal relationships, community functioning, social problem solving, and work performance. Additionally, social functioning affects the prognosis of schizophrenia patients. The symptoms of impaired social functioning are directly influenced by numerous factors, and it is thought that additional factors exert indirect influences. Although impairment in social functioning and psychopathology is well documented, there is no consensus as to which symptom of psychopathology is most closely related to social functioning. This study aims to develop an estimation algorithm for items on the Positive and Negative Syndrome Scale (PANSS) from the Specific Level of Functioning (SLOF) assessment scale.

Methods: Data were obtained from 328 subjects with schizophrenia or related disorders who participated in one of two clinical trials and completed both a PANSS and SLOF evaluation. Analysis was designed to establish the degree to which each item on the PANSS provides information about a subscale of social functioning (Interpersonal Relationship, Social Acceptability, Activities and Work Skills). The following analytic steps were performed: 1) a path analysis was done to identify which items of the PANSS links to each of the 4 domains; 2) each of the 30 items of the PANSS was used an independent variables into a data mining decision tree using discriminate analysis and regression analysis to select significant predictive items that influence each of the 4 domains of social functioning; 3) a non-parametric item response model (IRM) was fitted on the resulting items along the underlying latent continuum (social functioning).

Results: Standardized Regression Estimates showed the following items from the PANSS were linked to social functioning domains, with higher scores on PANSS items showing higher impairment in each domain: Social Acceptability (Delusions, 0.26; Excitement, 0.25; Grandiosity, 0.23; Blunted Affect, 0.20; Unusual Thought Content, 0.29), Interpersonal Relationship (Hostility, 0.21, Suspiciousness, 0.26; Passive Apathetic Social Withdrawal, 0.23; Poor Impulse Control, 0.23), Activities (Active Social Avoidance, 0.24; Disorientation, 0.26), and Work Skills (Hallucinatory Behavior, 0.30, Conceptual Disorganization, 0.25; Difficulty in Abstract Thinking, 0.25; Lack of Judgment/Insight, 0.30). Item Response Analysis shows high test information function (0.71) for the Social Acceptability and Interpersonal Relationship (0.73) domains.

Conclusions: This study examined a mapping algorithm suitable for linking scores on PANSS items to social functioning domains. Findings indicate that items from the PANSS can be used to estimate domains of social functioning, which can help guide researchers and clinicians in developing treatment strategies when formal social functioning measures are not available. These results may be used to predict social functioning, schizophrenia prognosis, and developing treatment interventions. These results also suggest that items on the PANSS can significantly predict social functioning in schizophrenia.

Disclosures: The authors have no disclosures.

19 Targeted Metabolomic Investigations as Potential Biomarkers in Mood and Anxiety Disorders:
Phenotypic Profiling to Clinical Trial Design

Szabo S,1,2, Kilts J,1,2, Naylor J,1,2, Strauss J,1,2, Morey R,1,2, Brancu M,1,2, Hamer R,3, Bradford D,1,2, VA Mid-Atlantic MIRECC Work Group1, Marx C1,2

1Durham Veterans Affairs Medical Center and VA Mid-Atlantic MIRECC, Durham, NC; 2Duke University Medical Center, Psychiatry and Behavioral Sciences; 3University of North Carolina School of Medicine, Departments of Psychiatry and Biostatistics, Chapel Hill, NC

Methodological Question Being Addressed: Can metabolomic profiles be incorporated into clinical trial design?

Introduction: Neurosteroids formed in the adrenal gland and de novo in the brain (i.e., dehydroepiandosterone) are altered in individuals with mood and anxiety disorders. While the relevance of serum constituents to neurobiologic function in psychiatric illness continues to be explored, the ability to characterize serum profiles linked to phenotypic characteristics in individuals with mood and anxiety disorders may aid clinical trial design.

Methods: Serum samples were obtained from a large cohort registry of 662 male Operation Enduring Freedom/Operation Iraqi Freedom (OEF/OIF) Veterans while capturing demographic and psychiatric symptom profiles. Samples were analyzed for levels of various neurosteroids in this cohort and metabolomic amino acid profiles (a smaller Veteran fraction of this cohort; n=90) with covariate analysis carried out for variables of relevance to mood and anxiety disorders.

Results: Resilience scores on the Connor-Davidson Resilience Scale (CD-RISC) positively correlated with dehydroepiandosterone sulfate (DHEAS) levels (r=0.15, p=0.0002), while global severity of symptoms (r=0.14, p=0.0001), anxiety (r=0.13, p=0.001), depression (r=0.13, p=0.001), and somatization (r=0.16, p=0.0001), assessed by the Symptom Checklist-90-R (SCL-90-R), were inversely correlated with DHEAS levels. DHEAS levels were significantly lower in Veterans with Davidson Trauma Scale (DTS) total scores ≥40 (consistent with PTSD; n=213) compared to Veterans with DTS total scores <10 (no/minimal PTSD symptoms; n=291), p=0.033. DHEAS levels were also significantly lower in Veterans with Beck Depression Inventory-II (BDI-II) total scores ≥20 (consistent with moderate depression; n=149) compared to Veterans with BDI-II total scores <10 (no/minimal depressive symptoms; n=359, p=0.026). Glycine, an excitatory amino acid and N-methyl-D-aspartate (NMDA) receptor modulator was significantly elevated in serum samples from Veterans reporting suicidal ideation (Beck Scale for Suicidal Ideation [BSS] score>0; n=19; p=0.043) as compared to Veterans reporting no suicidal ideation (BSS score=0, n=71). Aspartate/asparagine and serine, also excitatory neurotransmitters, were non-significantly increased in Veterans reporting suicidal ideation (p=0.097 and p=0.082, respectively). In contrast, arginine (nitric oxide [NO] precursor) and citrulline (byproduct of NO formation), were non-significantly decreased in Veterans reporting suicidal ideation (p=0.097 and p=0.093, respectively).

Discussion: DHEAS is a neurosteroid that is stable in serum that can potentially serve as a biomarker of resilience in Veterans with likely relevance in predisposition of developing depression and PTSD. As suicide is prevalent in these disorders, serum amino acid profiles in Veterans consistent with enhanced signaling and compromised NO formation may be a marker of risk. While understanding the impact of DHEA modulation of NMDA receptor and NO function in the brain may lead to pathophysiologic insights, serum levels of neurosteroids and amino acids may currently serve to function as fingerprints of disease. When applied to clinical trial design the potential for characterizing homogenous subject populations and treatment impact (i.e., benefits and side-effects) may lead to greater and early insights during clinical investigations of investigational compounds.

Disclosures: This work was funded by the VA Mid-Atlantic MIRECC (JAF), VA Mid-Level Career Development Transition Award (CEM). CEM is an applicant/co-applicant on pending patent applications for the use of neurosteroids in CNS disorders.

Examining for Potential Duplicate Patients in Clinical Trials: CATIE and STAR*D

Rabinowitz J, Rabinowitz Y

Bar Ilan University, Israel

Methodological Question Being Addressed: Determining potential duplicate patients within a clinical trial.

Introduction: Including patients who participate concurrently in more than one clinical trial, or who have recently participated in another trial, could severely bias trial results. There is mounting anecdotal evidence from people conducting clinical trials, but scarcely no published data from analyses of clinical trial data, that suggests that there are
people who are concomitantly multiply enrolled in clinical trials. Some have suggested that this is being accelerated by the emergence of websites to help people find trials, promote trial recruitment, flaunt the economic benefit and may be used to easily locate clinical trials.

**Aim:** Identify potential duplicate patients within CATIE and STAR*D.

**Methods:** Using the DupCheck algorithm, patients in the CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness) study (n=1460) matching on all of the following 8 criteria were identified: (1) age (date of birth not provided) within one year to account for possible second enrollment after birthday, (2) height, (3) BMI category (5 categories), (4) sex, (5) race (Caucasian yes/no), (6) Hispanic (yes/no), (7) Marital status (5 categories), (8) Level of education (8 categories). Matching in STAR*D (n=4042) was done using (1) age, (2) sex, (3) ethnicity (5 categories), (4) Hispanic (yes/no), (5) marital status (6 categories), (6) year of schooling and (7) residence (8 categories). Simulations were done to examine the likelihood that the above criteria would identify true duplicates.

**Results:** In CATIE 49 matches on the 8 criteria were found representing 87 study patients (29 had one match, seven had two matches and two had three). In STAR*D 103 matches on the 7 criteria were found representing 213 patients (96 had one match, 7 had three matches). The results of the simulations of 1000 trials with the sample size of CATIE, based on the distribution of CATIE subjects on the matching variables, suggests that the chance of incorrect positive identification of duplicates using this algorithm was less than 7% and the same for STAR*D less than 5%. Thus 3 of the matches in CATIE, and 5 of the matches in STAR*D, may have been false positives. This is based on false positive duplicate rate using these matching variables in samples of this size.

**Discussion:** Results suggest that like many trials, CATIE and STAR*D appear to have included duplicate patients. Having data on date of birth would have further reduced the chance of false positives. Based on population based birth cohort data, and simulations based on clinical trial populations, rate of false positives can be reduced to less than 1% by using complete date of birth, sex, initials and height. We only examined matches within each of these studies, leaving open the question of how many patients were concomitantly enrolled in other studies. We have established a free to use cross sponsor registry, DupCheck.org. of doubly encrypted patient demographic data to examine, prospectively and historically, the occurrence of duplicate patients using de-identified data. DupCheck.org allows screening out duplicate patients before enrolment and also re-analyzing completed trials after removing duplicate patients.


**Acknowledgments:** Data used were obtained from the limited access datasets distributed from the NIH-supported CATIE-Sz and STAR*D. CATIE was supported by NIMH Contract #N01MH90001 to the University of North Carolina at Chapel Hill. STAR*D was supported NIMH contract #N01MH90003 to the University of Texas Southwestern Medical Center.

**Disclosures:** First author may monetize DupCheck.org at some later time.

21 Superior Reproducibility of a “Mini-PANSS” Factor Model Across Samples: A Cross-Validation Factor Analytic Study of PANSS Factorial Resilience

Andersen A\(^1\), Chung H\(^4\), Li, Q\(^2\), Wilcox M\(^2\), Salvadore G\(^2\), Savitz A\(^2\), Nuamah F\(^2\), Riese S\(^3\), Bilder R\(^1\)

\(^{1}\)Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA \(^{2}\)Janssen Research and Development, \(^{3}\)Department of Psychology, University of California, Los Angeles, CA, \(^{4}\)Janssen Scientific Affairs

**Methodological Question Being Addressed:** Do PANSS factor models fail to replicate because of the underlying assumption that items fall into one, and only one, factor? Are restricted models more stable?

**Introduction (Aims):** Many published factor models of Schizophrenia empirically produced by exploratory factor analysis (EFA) fail to replicate across samples and related disorders; two recent works have failed to confirm 29 and 25 published models respectively on two different datasets. Using cross-validation, we test whether “mini-PANSS” factor models are more resilient on new data samples by proposing models with EFA and evaluating them on new data folds using confirmatory factor analysis (CFA).

**Methods:** Using 10-fold cross-validation, we propose new models using EFA and test these models with CFA on new samples. In order to improve model generalizability we varied item inclusion thresholds and allowed items to contribute to more than one factor (cross-loadings). Additionally, we compared four factor rotation methods to examine the influence of rotation on model fit, leading to a total of 400 models being created and tested on different sets of data.
The thresholds act as a gateway for items being included in a model. High thresholds can prevent items from contribution to more than one construct, resulting in sparse factor models. Moderate thresholds include all items usually but may result in items loading on only one factor. Low thresholds may include all items and allow items to influence more than one factor, thereby revealing inter-factor correlation.

**Results:** Our results suggest that PANSS factor models fail to replicate when they attempt to force every PANSS item into exactly one category. PANSS models with items removed were more resilient and reproduced better as the different random subgroups were tested. Rotation method did not significantly influence model resiliency. This suggests an “all or nothing” approach to item inclusion leads to superior models. The strongest PANSS models were those with high thresholds, which excluded 40% of the total PANSS items. Low-threshold models also performed well, and included all PANSS items with cross-loadings when the loadings were above the threshold. Moderate threshold items with 30 items which excluded cross-loadings performed most poorly.

**Conclusions:** The strong performance of sparse factor models suggests that low-loading items are sources of variability in the PANSS; these low-loading items need multiple measurements (i.e. cross-loadings) in order to be reproducible across samples. Given that factorial measurement variance has implications beyond the factor structure itself, it is important to identify how this can be reduced in the PANSS. It is possible that eliminating low-loading PANSS items would reduce measurement variability not just in the PANSS factor models, but also in the PANSS scale itself. Moreover, having a PANSS with limited items that are robust and have high loadings may ease the burden of administration of the scale and make it more practical for routine clinical work.

**Disclosures:** Dr. Chung is a full-time employee of Janssen Scientific Affairs. Drs. Li, Wilcox, Salvadore, Savitz, and Nuamah are full-time employees of Janssen Research and Development.

---

### Comprehensive Assessment of Functioning as an Endpoint in a Randomized Withdrawal Study

Turkoz I, Ascher S, Fu D-J, Nuamah I, Olson W, Alphs L  

1Janssen Research & Development, LLC, Titusville, NJ, 2Janssen Scientific Affairs, LLC, Titusville, NJ

**Methodological Question Being Addressed:** How does one evaluate the totality of the evidence supporting maintenance of functioning, as measured by the Personal and Social Performance (PSP) scale, in long-term, placebo-controlled, relapse prevention trials or outcomes trials after stabilization?

**Introduction:** In a 15-month relapse prevention trial of 334 randomized subjects (1:1 active treatment vs. placebo), functionality outcome data were obtained monthly. 51% of patients discontinued before the 15-month endpoint visit because of relapse, withdrawal of consent, loss to follow-up, and/or adverse events. In addition to substantial missing data at the study endpoint, the ability to validly detect change in the PSP scale score was limited by the ethical need to intervene as early as possible before a full relapse with attendant full deterioration in functioning occurred. Early intervention to meet these ethical goals had the effect of significantly reducing the magnitude of changes in the PSP, thereby minimizing the potential to detect differences in functioning between the two treatment arms. Indeed, more than one-fourth of relapse events identified in this study represented early symptom deterioration before hospitalization occurred. The objective of this analysis is to provide a comprehensive assessment of the PSP scale in estimating treatment differences in a relapse prevention trial. These illustrations allow us to compare results using selected statistical methods to assess consistency and robustness of PSP findings at study endpoint.

**Methods:** Three statistical approaches analyzing PSP as three types of variables were carried out to assess the robustness and consistency of findings in patient functioning at study endpoint: (1) methods that examine the PSP score as a continuous variable, (2) those that examine the PSP as a categorical variable, and (3) those that examine time to clinically significant decrements in PSP. Each statistical approach provides results and interpretations that contribute to the totality of evidence in estimating treatment differences. The mixed model repeated measures (MMRM) approach, analyzing PSP as a continuous endpoint, was prespecified as the primary analysis. Supplementary analytical methods were carried out within each of the three statistical approaches to provide further evidence to confirm and compare study results. To assess the findings and the assumptions of MMRM analysis when examining the PSP as a continuous variable, three additional methods were implemented: (1) tipping point analysis, (2) multiple imputation, and (3) pattern mixture models.

**Results:** The comparison between the two treatment groups for mean change in PSP score from baseline at month 15 was significant (p=0.014) using the MMRM approach. The LS-mean (95% CI) difference between treatment groups in change scores at month 15 was 3.3 (0.68, 5.95) in maintaining function, favoring the active treatment group.
A PSP score between 71 and 100 indicates good functioning, a score between 31 and 70 signifies varying degrees of difficulty, and a score ≤30 indicates such poor functioning that subjects require intensive supervision. The proportion of subjects with good functioning decreased from 50.6% at DB baseline to 41.1% at the 15-month DB endpoint in the placebo group, whereas subjects in the active treatment group maintained good functioning with 57.9% at DB baseline and 59.9% at DB endpoint (p<0.01).

A decrement of at least 7 points on the PSP is considered a meaningful, clinically relevant worsening of patient global functioning. Times to 7- and 10-point decrements in PSP scores were evaluated. Between-treatment group differences favored the active treatment (p<0.01) for time to both 7-point and 10-point decrements.

Conclusion: The totality of evidence presented here shows that the active treatment group is superior to placebo in maintaining functioning as measured by the PSP in a relapse prevention trial.

Support: Janssen Scientific Affairs, LLC

23 How Does the NSA-4 Compare to the NSA-16?

Williams, J1,2, Garzio L1, Osman D1, Popp D1

1MedAvante, Inc., 2Columbia University Departments of Psychiatry and Neurology

Methodological Question Being Addressed: Do the psychometric properties of the NSA-4 justify its use as a substitute for the full NSA-16 scale in clinical trials?

Introduction: The 16-item Negative Symptom Assessment (NSA-16) is increasingly used as a validated measure to track response to treatment of negative symptoms in clinical trials of schizophrenia. The NSA-16 takes up to a half hour to administer. Alphs et al. (2011) have proposed a four-item version, the NSA-4, as a reliable and valid brief alternative. The current study is an attempt to replicate the previous findings of Alphs et al. in two randomized clinical trials.

Methods: Central raters from a well-trained independent and blinded cohort with ongoing monitoring to ensure calibration and prevent drift interviewed subjects in two placebo-controlled randomized double-blind studies of schizophrenia with prominent negative symptoms using live two-way videoconferencing at screen, baseline, and 11 more visits, including end point. At each visit, raters administered the PANSS immediately followed by the NSA-16. Correlations between the NSA-16 and the NSA-4 were examined for the NSA global rating, the PANSS negative and positive subscales, and several Marder factors. In addition, Cronbach’s alpha and interrater reliability were examined for both scales.

Results: The NSA-16 was administered 2804 times by 29 central raters to a total of 483 subjects in the two trials. Overall, the correlation between the total scores of the NSA-4 and NSA-16 was very good (0.86). Good convergent validity of the NSA-4 was demonstrated by correlations between the NSA-4 and the NSA global rating, as well as the PANSS negative subscale and the PANSS negative symptoms Marder factor of 0.67, 0.73, and 0.73, respectively. Alphs et al. found these correlations to be 0.68, 0.52, and 0.57. Divergent validity in our sample was demonstrated by low correlations between the NSA-4 and the following PANSS Marder factors: anxiety/depression, disorganized thought, hostility/excitement, and PANSS positive symptoms: -0.11, 0.29, 0.03, and 0.13, respectively. Comparable values in the Alphs et al. study were: -0.03, 0.42, 0.06, and 0.23.

Cronbach’s alpha, as expected for a shorter scale, was lower for the NSA-4 in our studies as well as the Alphs et al. study. We found it to be 0.65, compared to 0.64 for Alphs et al.; the NSA-16 in our study was 0.87 compared to 0.85 for Alphs et al. Our interrater reliability estimate for the NSA-16 was 0.97, compared to Alphs et al. 0.87. Our ICC for the NSA-4 was 0.94 while Alphs et al. was 0.82, both in the excellent range.

Conclusions: The PANSS and NSA-16 in this study were not administered independently of one another, so the usefulness of the NSA-4 alone can only be evaluated in the context of its pairing with the PANSS. Overall, results were very similar to those obtained by Alphs et al. In the hands of highly trained and calibrated central raters, the NSA-4 had very good overall agreement with the NSA-16, and even higher convergent and divergent validity and interrater reliability than was demonstrated by Alphs et al.


Disclosure: Authors are all employees of MedAvante, Inc.
**How Duplicate Subjects Impact Clinical Trial Data: Under-compliant, Over-responding and Statistically Overlooked**

Shiovitz T\(^1\), Zarrow M\(^1\), Mangano T\(^2\)

\(^1\)California Neuroscience Research, Sherman Oaks, CA; \(^2\)CTSdatabase, LLC, Beverly Hills, CA

**Methodological Question Being Addressed:** How duplicate subjects impact clinical trial data.

**Introduction:** Duplicate and professional subjects are of growing concern in the conduct of clinical trials. They represent a subset of problematic study subjects which include (and overlap with) non-compliant subjects, placebo responders and subjects who do not reflect the patient population. Duplicate and professional subjects may alter their symptoms or their medical histories in order to participate in multiple studies and collect stipends. They may participate in studies for different indications and are will to go to great distances to participate in studies at different sites. We will explore how the participation of duplicate subjects translates into failed studies.

**Methods:** To quantify the duplicate subject problem, numbers and types of duplicate subjects were obtained from the CTSdatabase subject registry. This data includes dates, locations and presenting indications of 6117 potential and consented US clinical trial subjects (predominantly in CNS indications such as Schizophrenia, Bipolar or Major Depression) who were entered into CTSdatabase between 10/31/2011 and 12/31/2013. Only subjects that signed an IRB-approved authorization and matched enough partial identifiers to be a virtually certain match (i.e. < 1 \( \times 10^{-7} \) likely to be a chance match with an existing database entry) were included. This degree of certainty is far higher than matching only date of birth, for example, which occurs with an age-dependent approximate frequency range of from 1 in 5,000 to 1 in 50,000. A literature search was also performed.

**Results:** 12.4% of subjects entered into CTSdatabase were virtually certain matches with subjects at another site within 180 days, 6.0% within 60 days and 3.7% within 30 days of presenting to the original site. Prescreens who were not entered into studies potentially overestimated the number of duplicates found. Conversely, duplicate subjects were underestimated because not all sites and sponsors participated in the registry. Overall, this data is consistent with reports that at least 10% of certain CNS subjects in the US may be duplicates. These duplicate subjects may represent many of the approximately 15-30% of subjects that are found by pharmacokinetic sampling to be noncompliant. Even a small change in the number of non-compliant subjects can require a substantial increase in sample size in order to maintain study power.

There is also evidence to suggest that duplicate subjects may appear compliant by pill count but do not take study medication. Examples of how non-compliant subjects respond like placebo subjects and how duplicate subjects may dramatically placebo respond have also been reported.

**Conclusion:** Duplicate subjects are a significant problem in clinical trials and may represent upwards of 10% of subjects in some studies. They are more likely to placebo respond, not take study medication and, if not accounted for in calculations of sample size, could lead to study failure. However, these subjects could be taken into account in statistical calculations and study design. An awareness of duplicate enrollment when calculating sample size, planning recruitment and choosing serial PK sampling or a newer compliance technology is critical. Use of a subject registry prior to randomization is a highly efficient way to reduce numbers of duplicate subjects entering studies. Unlike potential problems with eliminating other non-compliant subjects from the ITT sample, duplicate subjects do not represent the non-compliance of real-world patients and therefore should be eliminated before randomization.

**Disclosure:** Dr. Shiovitz has ownership interest in CTSdatabase, LLC. None of the authors received financial support.

**Recruitment Challenges for Negative Symptom Studies of Schizophrenia**

Dholakia S, Maples N, Brown M, Velligan D

*The University of Texas Health Science Center at San Antonio, TX*

**Methodological Question Being Addressed:** What is the feasibility of using strict criteria for studies on negative symptoms in Schizophrenia?

**Introduction:** Negative symptoms are widely accepted as a component of schizophrenia for which there is currently no adequate pharmacological or therapeutic treatment. Recently, there has been increasing emphasis on the development of both pharmacologic and psychosocial treatments to address negative symptoms. However, there is a lack of consensus and standardization on what criteria should be used to select participants for negative symptom studies.
Methods: We describe recent recruitment challenges and issues in a psychosocial treatment study addressing persistent (≥6 months), primary (not due to depression, positive symptoms, extrapyramidal side effects), negative symptoms. Recruitment was conducted at community behavioral health clinics in Austin and San Antonio, TX. Recruitment was based on chart review as well as initial screening by recruiters.

Results: Early recruitment lasted for 7 months and included an initial prescreening, including an evaluation of negative symptom stability, completed by recruiting staff prior to the full assessment. 49 individuals with at least moderate negative symptoms in 2 of 6 domains on the Scale for the Assessment of Negative Symptoms (SANS) completed a baseline assessment. Of these, 10 were disqualified on the basis of positive symptom severity and 3 on the basis of depressive symptomatology. Thus, 26% of recruits did not meet baseline criteria, even with the initial pre-screening. Positive symptom severity (exclusion) was established with Brief Psychiatric Rating Scale (BPRS) scores ≤5 on items measuring delusions (unusual thought content) and hallucinations. For inclusion, depression severity was required to be ≤3 on the depression item.

Due to difficulties recruiting individuals who met study criteria, the positive symptom and depression criteria were relaxed. Positive symptoms severity was exclusionary only when they would interfere with the treatment. Depression symptomatology was replaced by suicidal ideation as exclusion criteria. The second phase of recruitment lasted for 9 months and included 42 individuals, 35 of whom completed a baseline assessment. Of these, 2 were disqualified on the basis of positive symptom severity and 2 on the basis of suicidal ideation. For various reasons, 7 individuals were recruited but did not complete baseline assessment. Over the course of the study, 2 subjects were disqualified due to medication side effects as established by a global score of 4 or higher on the Extrapyramidal Symptom Rating Scale (ESRS), and 4 subjects were lost to follow-up or no longer interested in study participation.

Conclusion: Given that free psychosocial treatment was offered and medication changes were not required, levels of non-eligibility after baseline may be acceptable for trials of psychosocial treatments. However, for pharmaceutical trials which have additional requirements, restrictive criteria may not be feasible.

Disclosure: The authors report no conflicts of interest for this work.

26 Risk Based Data Quality Monitoring Utilizing Data Analytics and Recorded PANSS Interviews in Global Schizophrenia Trials

Daniel D, Kott A

Bracket Global, LLC

Methodological Question Being Addressed: Can a battery of data quality markers effectively detect poor performing sites?

Introduction: Outlier analysis of blinded data for aberrant rating patterns and patient selection anomalies can be paired with audio/video surveillance to cost-effectively identify at-risk sites in global schizophrenia clinical trials.

Methods: Utilizing centralized, blinded data quality monitoring of 41,555 PANSS assessments in ten international schizophrenia clinical trials, norms were created for data patterns selected by sponsors as potentially at risk for measurement error or idiosyncratic patient selection. Based on these risk factors, a composite score or “dashboard” was created ranking each site based on quality measures. Sites of concern were subsequently subjected to more intensive, remote, centralized review of recorded patient interviews by external experts. The quality of recorded interviews and ratings was remotely assessed by independent reviewers for 2,943 PANSS assessments.

Results: Based on independent review of audio and/or video recorded PANSS assessments, interview quality was rated as excellent, adequate with some deficiencies or inadequate in 75.44 % (n=2221), 23.2% (n=683) and 1.36% (n=40) of visits, respectively. Proper application of the PANSS instructions and anchor points was independently rated as excellent, adequate with some deficiencies, or inadequate in 75.98 % (n=2221), 22.8% (n=671) and 1.22% (n=36) of visits, respectively.

The following examples illustrate adaptive monitoring. Sites 397 and 762 were evaluated on three risk factors specified by the sponsor: 1) large between-visit PANSS changes; 2) erratic PANSS changes; and 3) 100% identical PANSS scores from visit to visit. If anomalies were determined by blinded data monitoring additional scrutiny was employed by external review of recorded patient visits.

Site 397 was an outlier on factors 1 and 2 (> 3 SD above the mean) but refused to allow interviews to be recorded for external review to allow independent assessment of measurement error.

Site 762 was not an outlier on large score or erratic score changes but more than 15% of visits were 100% identical.
Recordings of patient interviews were scrutinized. The proportion of discordant PANSS ratings (>2 difference between site and independent rater) exceeded 60%.

Conclusion: Risk based outlier analysis of blinded data for aberrant rating patterns and patient selection anomalies can be paired with audio/video surveillance to cost effectively identify at risk sites in global schizophrenia clinical trials. Allowing sites to “opt out” of audio/video surveillance complicates interpretation of data anomalies. In addition, audio/visual surveillance has the potential to identify endpoint scoring irregularities that may not emerge in outlier analysis.

27 Evaluation of Daytime Sleepiness in Patients with Schizophrenia Treated with Atypical Antipsychotics: Results from a Randomized, Double-blind, Placebo-controlled Trial

Loebel A1, Siu C2, Cucchiaro J1, Pikalov A1, Harvey P3

1Sunovion Pharmaceuticals Inc., Marlborough, MA, 2Data Power (DP), Inc., Flemington, NJ, 3University of Miami Miller School of Medicine, Miami, FL

Methodological Question Being Addressed: 1) How daytime sleepiness (also known as daytime somnolence or sedation) associated with antipsychotic treatment can be evaluated in a clinical trial involving patients with schizophrenia, and 2) How daytime sleepiness might impact treatment outcomes.

Introduction: Daytime sleepiness (also known as daytime somnolence or sedation) associated with antipsychotic treatment may adversely impact treatment outcomes and functional performance and quality of life. The aim of this post-hoc analysis was to compare the effects of 2 atypical antipsychotic agents, lurasidone (80 mg/d or 160 mg/d) and quetiapine XR (600 mg/d), on daytime alertness, and to evaluate the effects of daytime sleepiness on treatment outcomes in patients with an acute exacerbation of schizophrenia.

Methods: Patients who met DSM-IV-TR criteria for schizophrenia were randomized to 6-weeks of double-blind treatment with fixed doses of lurasidone 80 mg/d (n=125), lurasidone 160 mg/d (n=121), quetiapine XR 600 mg/d (n=119), or placebo (n=121), all dosed once-daily in the evening, with food. Daytime sleepiness was assessed using the Epworth Sleepiness Scale (ESS) at baseline, week 3, and week 6 visits. The ESS is a patient-reported questionnaire, which consists of 8 items that describe various routine daily life situations. Subjects rate the likelihood of dozing or falling asleep (distinguishing dozing behavior from feelings of tiredness) in each of these situations, on a 4-point scale from 0 to 3. It has been validated in both case-control studies of normal subjects and several patient populations with different types of sleep disorders. In addition, the ESS has been utilized for the assessment of daytime sleepiness in patients with schizophrenia. A mediation regression approach was applied to explore the potential effect of daytime sleepiness (mediator) and its association with study antipsychotic treatments (exogenous causal variable) on changes in agitation (as assessed by the PANSS-EC), cognitive (assessed by CogState Computerized Schizophrenia Battery), and functional capacity (assessed by the UPSA-B total score).

Results: Daytime sleepiness improved in the lurasidone and placebo-treated groups but worsened in the quetiapine XR treatment group when compared to placebo (p=0.001) and to either dose of lurasidone (both p<0.01). Sedation associated with quetiapine XR treatment mediated an improvement in agitation and a worsening in functional capacity; these mediating relationships were not observed for the lurasidone or placebo treatment groups.

Conclusions: In this 6-week double-blind study, treatment with lurasidone 80 mg or 160 mg, administered once-daily in the evening, was associated with a reduction in daytime sleepiness similar in magnitude to placebo, while quetiapine XR 600 mg/d was associated with a significant increase in daytime sleepiness, compared to both lurasidone dose groups and placebo. Daytime sleepiness was associated with improvement in agitation and worsening in functional capacity for quetiapine XR, but not lurasidone or placebo-treated patients. Our findings suggest that daytime somnolence may have a significant impact on cognitive and functional outcomes. In addition, these findings support the use of the ESS as a brief and practical tool for the assessment of daytime sleepiness in patients with schizophrenia and provide evidence for its sensitivity to change during treatment with antipsychotic agents.

Disclosures: Drs. Cucchiaro, Loebel and Pikalov are employees of Sunovion Pharmaceuticals Inc. Dr. Siu has received payment for consulting from Pfizer Inc., Sunovion Pharmaceuticals Inc., and Takeda. Dr. Harvey serves as a consultant/advisory board member for Abbvie, Boehringer Ingelheim, Bristol-Myers-Squibb, Forest Labs, Genentech, , Roche, Shire, Sunovion, and Takeda.
Comparative Sensitivity of Three Efficacy Measures for Detecting Treatment Effects in a Short-term, Double-blind, Placebo-controlled Study in Bipolar Depression

Pikalov A1, Tsai J1, Kroger H1, Cucchiaro J1, Loebel A1

1 Sunovion Pharmaceuticals Inc., Marlborough, MA and Ft. Lee, NJ

Methodological Question Being Addressed: Is there a differential advantage for global vs. symptom-based scales, or clinician vs. patient-rated scales in detecting improvement in patients enrolled in a short-term bipolar depression trial?

Introduction: The aim of this post hoc analysis was to evaluate the comparative sensitivity of three standard efficacy measures for detecting significant improvement in depressive symptoms in patients enrolled in a short-term study of lurasidone in bipolar I depression.

Methods: Patients with bipolar I depression were randomized to 6 weeks of double-blind monotherapy with lurasidone (20-60 mg/d; N=161 or 80-120 mg/d; N=162) compared with placebo (N=162). Outcomes of the following efficacy measures were collected weekly: (1) the 10-item clinician-rated MADRS total scale; (2) the single-item clinician-rated Clinical Global Impression, Bipolar Severity of Depression (CGI-BP-S) score; and (3) the 16-item patient-rated Quick Inventory of Depressive Symptomatology - Self Report (QIDS-SR16). In addition, we examined results from the 6-item clinician-rated MADRS “core” depression subscale (Bech et al, Psychopharm [Berlin] 2002;163:20-25).

Results: The table below summarizes efficacy outcomes at Weeks 2 and 6.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>LS mean change at Week 2</th>
<th>P value (Effect size)</th>
<th>LS mean change at Week 6</th>
<th>P value (Effect size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MADRS total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lurasidone 20-60 mg</td>
<td>161</td>
<td>-7.3</td>
<td>0.040 (0.24)</td>
<td>-15.4</td>
<td>&lt;0.001 (0.51)</td>
</tr>
<tr>
<td>Lurasidone 80-120 mg</td>
<td>162</td>
<td>-8.1</td>
<td>0.002 (0.36)</td>
<td>-15.4</td>
<td>&lt;0.001 (0.51)</td>
</tr>
<tr>
<td>Placebo</td>
<td>162</td>
<td>-5.6</td>
<td></td>
<td>-10.7</td>
<td></td>
</tr>
<tr>
<td>MADRS 6 (Core)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lurasidone 20-60 mg</td>
<td>161</td>
<td>-4.9</td>
<td>0.017 (0.27)</td>
<td>-10.4</td>
<td>&lt;0.001 (0.54)</td>
</tr>
<tr>
<td>Lurasidone 80-120 mg</td>
<td>162</td>
<td>-5.4</td>
<td>0.001 (0.37)</td>
<td>-10.4</td>
<td>&lt;0.001 (0.54)</td>
</tr>
<tr>
<td>Placebo</td>
<td>162</td>
<td>-3.6</td>
<td></td>
<td>-6.9</td>
<td></td>
</tr>
<tr>
<td>CGI-BP-S Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lurasidone 20-60 mg</td>
<td>161</td>
<td>-0.71</td>
<td>0.009 (0.30)</td>
<td>-1.83</td>
<td>&lt;0.001 (0.61)</td>
</tr>
<tr>
<td>Lurasidone 80-120 mg</td>
<td>162</td>
<td>-0.77</td>
<td>&lt;0.001 (0.37)</td>
<td>-1.71</td>
<td>&lt;0.001 (0.50)</td>
</tr>
<tr>
<td>Placebo</td>
<td>162</td>
<td>-0.46</td>
<td></td>
<td>-1.14</td>
<td></td>
</tr>
<tr>
<td>QIDS-SR16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lurasidone 20-60 mg</td>
<td>157</td>
<td>-3.8</td>
<td>0.036 (0.25)</td>
<td>-6.9</td>
<td>0.002 (0.39)</td>
</tr>
<tr>
<td>Lurasidone 80-120 mg</td>
<td>160</td>
<td>-4.0</td>
<td>0.008 (0.31)</td>
<td>-7.5</td>
<td>0.001 (0.53)</td>
</tr>
<tr>
<td>Placebo</td>
<td>160</td>
<td>-2.9</td>
<td></td>
<td>-5.2</td>
<td></td>
</tr>
</tbody>
</table>

Conclusions: Comparative sensitivity was observed among all evaluated measures. The effect size data indicate that a single item global severity measure (CGI-BP-S) performs as well as the MADRS total and MADRS-6 subscale for detecting both early improvement (Week 2), and improvement at Week 6. The performance of the patient-rated QIDS-SR16 was comparable to the MADRS total score for detecting significant improvement in symptoms of depression.

Disclosures: Sponsored by Sunovion Pharmaceuticals Inc. (NCT00868452)

Concordance Between Clinician and Centralized Computer Ratings in a Bipolar Maintenance Study of ELND005 (Scyllo-inositol)
Kurth M1, Crans G1, Bairu M1, Sachs G2, Abushakra S1

1Elan Pharmaceuticals, Cambridge, MA, 2Department of Psychiatry, Massachusetts General Hospital and Bracket Global, Boston, MA

Methodological Question Being Addressed: To describe concordance in depression and mania ratings performed by clinical investigators and those by a centralized computer system (Computerized Rater Station, Reilly-Harrington et al. 2010, Sachs et al. 2012), and to analyze which ratings may better predict relapses of mood episodes in a bipolar type I (BPD) maintenance study.

Introduction: Elevations of brain myo-inositol are described in BPD patients with mania or depression (Davanzo et al. 2001, Moore et al. 1999). Myo-inositol dysregulation is likely associated with abnormal neuronal signaling in limbic frontal networks; the mood stabilizing effect of lithium is thought to be mediated by myo-inositol regulation (Moore et al. 1999). In AD patients, ELND005 (Scyllo-inositol) has shown beneficial trends on cognition and depression/anxiety (Study AD201, Salloway et al. 2011, Abushakra et al. 2012) and lowered myo-inositol brain imaging (Tariot et al. 2012). ELND005 is therefore being evaluated as a potential maintenance treatment for BPD (Study BPD201, clinicaltrials.gov NCT01674010). Concordance rates between clinician and centralized computer ratings for depression (MADRS) and mania (YMRS) scores are analyzed across all visits, and compared in the groups who developed a mood episode (recurrence) to those without recurrence in the open phase of the ongoing study.

Methods: Study BPD201 enrolls euthymic BPD patients (MADRS and YMRS ≤ 12, ages 18-65) who had experienced a mood episode in prior 4 months, and who were successfully treated with standard of care medications. Subjects enroll in a 16 week open treatment phase with ELND005 (500mg BID), while their other medications are gradually withdrawn except for either valproate or lamotrigine. Subjects who remain euthymic at 16 weeks (both scores ≤ 16) are enrolled in a placebo-controlled phase for up to 48 weeks. The primary outcome measure is time to recurrence of any mood episode. MADRS and YMRS are performed at all clinic visits by both clinician and by Computerized Rater Station (Reilly-Harrington et al. 2010, Sachs et al. 2012). Concordance was assessed using weighted Kappa statistic.

Results: Among the first 210 enrolled subjects, the mean age was 45 years (51% female; 67% Caucasian, 23% African American, 7% Hispanic). In the open treatment phase, a total of 77 completed week 16 and entered the randomized phase, and 23 had mood episode (29% recurrence rate). Baseline mean scores for computer/clinician ratings were MADRS: 4.8/4.0 and YMRS: 4.0/3.2. Across all open phase visits, mean difference between computer/clinician MADRS and YMRS were 0.65 and 0.96 (n= 1720 and 1714), with MADRS computer ratings higher in 33% and lower in 19%; YMRS ratings were higher in 48% and lower in 12%. Baseline concordance rates in all subjects, those with recurrence, and without recurrence for MADRS and YMRS were: 0.71 and 0.65, 0.59 and 0.59, and 0.72 and 0.66 respectively. Across all visits, MADRS and YMRS concordance rates were: 0.78 and 0.66 (all), 0.78 and 0.72 for (with recurrence), and 0.77 and 0.65 (without recurrence).

Conclusions: Concordance rates between clinician and computer ratings for depression and mania are good in this bipolar maintenance study. Clinician/computer concordance is consistently higher for MADRS than YMRS scores regardless of mood episode recurrence. Computer YMRS ratings are higher than clinician ratings in almost half of visits, but the mean difference is small (< 1 point). YMRS concordance rates are highest among subjects with mood episode recurrence.

Disclosures: Drs. Kurth, Crans, Bairu and Abushakra are full time Elan employees, and stockholders in Elan Pharmaceuticals; Dr. Sachs is an employee of Bracket Global and is a consultant for Elan Pharmaceuticals for this study.

Drug Development for Agitation and Aggression in Alzheimer’s Disease (AD): Study Design and Outcome Measures in Phase 2 Study of Scyllo-inositol (ELND005)

Abushakra S1, Kesslak JP1, Kurth M1, Fan F1, Bairu M1, Miller D2, Lyketsos C3

1Elan Pharmaceuticals, Cambridge MA, 2Bracket Global, Wayne PA, 3Department of Psychiatry, Johns Hopkins University, Baltimore MD

Methodological Question Being Addressed: Drug development for Agitation and Aggression in Alzheimer’s Disease requires the establishment of widely accepted diagnostic criteria and appropriate behavioral outcome measures.

Introduction: Agitation and Aggression are among the most disruptive of neuropsychiatric symptoms (NPS) in AD dementia. Agitation and Aggression occur in up to 50% of AD patients, are associated with increased morbidity, caregiver burden and healthcare cost. Few drugs have shown definitive efficacy and safety in these patients. To date, there are no FDA approved drugs for this indication. Agitation and Aggression are likely related to monoaminergic...
imbalance and synaptic dysfunction in cortical networks (Lyketsos, 2006), possibly due to amyloid toxicity and disruption of neuronal signaling. In AD patients, ELND005 (Scyllo-inositol) has shown amyloid lowering effects in CSF (Salloway et al. 2011), myo-inositol lowering effects on brain imaging (Tariot et al. 2012), and beneficial trends on cognition and agitation (Abushakra et al., 2012). ELND005 is being evaluated as a potential treatment for Agitation and Aggression in AD (clinicaltrials #NCT01735630). We herein describe the operational diagnostic criteria and novel outcome measures utilized in this trial.

**Methods:** Study ELND005-AG201 is a 12-week, 2 arm study (1 ELND005 arm and placebo), in AD patients ages 50-88 with dementia (MMSE range 8-23). Subjects are required to have no active or untreated underlying metabolic, infectious, or medical condition causing their NPS, and to be on stable doses of AD drugs and/or antipsychotics (if any). The 12-item Neuropsychiatric Inventory (NPI, Cummings et al. 1999) is used for screening. Subjects are eligible if screening and baseline NPI-agitation/aggression score is ≥ 4 (at least moderate severity). The primary outcome measure is the summed agitation and aggression scores from the NPI-C, which separates Agitation and Aggression into distinct domains (expanded and clinician rated NPI, De Medeiros et al. 2010). Clinician global impression of change (ADCS-CGIC) for agitation/aggression is a key secondary measure to support clinical meaningfulness of NPI-C treatment effects.

**Results:** Of the first 100 patients enrolled, mean age was 75 years (55% female; 89% Caucasian and 9% African American). Among randomized subjects, baseline mean MMSE was 15.4 (range=6—26); NPI-total score was 47.6 (5—119), and NPI Agitation/Aggression scores was 7.2 (4—12). The NPI-C Agitation and Aggression scores were, respectively 12.7 and 5.5 (0—30 and 0—21; SD: 6.5 and 4.6). Baseline scores of other NPS (depression, anxiety, apathy) and distribution of NPI-C agitation and aggression scores amongst mild, moderate and severe AD subgroups will be described.

**Conclusions:** The operational diagnostic criteria in this study, as applied by clinical investigators, identify a population of AD patients with at least moderate levels of agitation/aggression requiring pharmacologic intervention. These criteria are similar to those recently proposed by the International Psychogeriatric Association task force. Performance of the NPI-C as outcome measure in this global study (English, Spanish, French versions) will inform its utility in future registration trials.

**Disclosures:** Drs. Abushakra, Kesslak, Kurth, Fan, and Bairu are full time Elan employees, and stockholders in Elan Pharmaceuticals; Dr. Miller is an employee of Bracket Global. Dr. Lyketsos is a consultant for Elan for this study.

### 31 Promoting Participation in Large, Multi-site Trials for Patients with Pediatric Cancer


*Children’s National Medical Center, Washington, DC, UT Southwestern Medical Center, Dallas, TX, CogState, Inc. and Yale Child Study Center, New Haven, CT, University of New Mexico Health Sciences Center, Albuquerque, NM, University of Texas Health Science Center, San Antonio, TX, University of Minnesota, Minneapolis, MN, United States Army Medical Research and Materiel Command, Fort Detrick, MD, University of Pittsburgh Medical Center, Pittsburgh, PA*

**The Methodological Question Being Addressed:** Can rates of participation and retention of children in neurocognitive testing during cancer trials be improved with the use of a brief computerized neurocognitive battery?

**Introduction:** Survivors of childhood acute lymphoblastic leukemia (ALL) are at risk for neurocognitive deficits, which are known to be associated with impaired school performance and can adversely impact QoL. Thus, use of neurocognitive outcomes as an endpoint to help identify patients for early intervention would be of substantial value, to help differentiate the risk-benefit profiles between treatments would arguably be of substantial value. Within the Children’s Oncology Group (COG), two strategies have emerged. One uses a brief (1 hour) neuropsychological monitoring battery (ALTE07C1) and the other uses computerized strategies. Leadership within COG selected a brief computerized battery (Cogstate) comprised of measures of processing speed, attention, memory, working memory and executive function to prospectively examine neurocognitive functioning in children with high-risk ALL in the context of a COG clinical trial. Here we report on interim analyses to examine the success of this approach in achieving three prespecified metrics: 1) at least 100 sites successfully participating, 2) recruiting at least 55% of eligible patients, and 3) achieving a data collection rate of 90%.

**Methods:** The sample includes English or Spanish-speaking children, aged 6-11 years, enrolled on a COG phase III ALL trial for newly diagnosed patients. Serial assessments are conducted every 6 months from 3 months post-diagnosis to 1-year post-treatment. Children complete a 25-minute computerized battery that can be administered by any professional in the oncology clinic. Data are automatically scored and uploaded creating an auditable trail of secure data for centralized data management.
**Results:** 115 sites are participating in the neurocognitive study and 126 of 170 eligible participants (70%) have been enrolled. For the first time point, 94% of participants have complete data, and 97% have complete data for those reaching the second time point. Only two children have been unable to complete testing within the data window. Data integrity checks, designed to detect poor effort or misunderstanding of the tasks, indicate that the majority of participants (>94%) provided valid data.

**Conclusions:** Data suggest that pediatric versions of Cogstate tasks can be used to conduct research assessments in large cancer trials in children. Use of a computerized battery may boost rates of accrual and adherence in multi-site clinical trials. Relative to even abbreviated (ALTE07C1) neurocognitive batteries, the scope and depth of data ascertained are somewhat limited by this approach; however, it may be advantageous when large, diverse samples are desired.

**Disclosures:** Dr. Harel is a full time employee of Cogstate, a cognitive test company that provided the tests used in this study.

### 32 High Fidelity: What is the Real Impact of a Data Monitoring Program on Data Quality?

Yavorsky C, DiClemente G, Wolanski W, Burger F  
*Cronos CCS, Lambertville, NJ*

**Introduction:** Risk-based data-monitoring is widely used in clinical trials to help manage risk implicit in subjectively derived outcome measures. Few studies to date have critically examined the impact of such programs. In this study our intention was to determine how well the method works and then to estimate the potential impact of non-intervention on statistical power.

**Methods:** The present study investigated a sample of subjects enrolled in a completed schizophrenia trial using the PANSS as the primary outcome measure. The data was processed daily and, if risks to data quality were detected, contact was initiated. Feedback was provided as necessary when problems in scale use were identified. The sensitivity and specificity were computed to determine the proportion of data correctly identified as problematic. A forward analysis estimating the impact of non-intervention with raters contributing poor quality data was conducted.

**Results:** The first analysis allowed for summarization of the overall efficacy of the method while identifying those raters whose error would have adversely impacted trial outcomes was used to re-estimate sample size. We theorized a non-intervention scenario with patients continuing to be assessed incorrectly and subtracting these. Using this estimated reduction in sample size we recalculated power based on original parameters and found a reduction from .90 to .78.

**Conclusions:** Risk-based data-monitoring can detect error within reasonable estimates and be addressed in-study. This analysis had two essential aims; to outline what the baseline rater error rate was in a typical study; and to determine the potential impact of a reduced sample size if this data containing error is excluded.

**Disclosure:** The authors are employed by Cronos CCS, and report no additional conflicts of interest.

### 33 Working Group Reports