ISCTM 12th Annual Scientific Meeting

16-18 February 2016

The Fairmont, Washington, DC

Poster Abstracts

Groupings:
Psychosis: 1-11
Other: 12-22
Methodology: 23-30

1 The Metabolism of a Psychosocial Treatment Intervention in Schizophrenia: The Helpful Habit Program

Docherty J,[1] Hatch A,[1] Profit D,[1] Lawson E,[1] Adenwala A,[2] Velligan D[3]

[1]Otsuka Pharmaceutical Development & Commercialization, Inc., Princeton, NJ, [2]GfK Custom Research LLC, Oakbrook Terrace, IL, [3]University of Texas Health Science Center, San Antonio, TX

Methodological Question Being Addressed: Substantial advance has been made in the standardization of what the psychotherapist does in a clinical trial. However, little is known about how that intervention is understood and digested (metabolized) by the patient. The question addressed here is: How are psychosocial treatments metabolized by patients and health care providers (HCPs) and how can this inform program refinement?

Introduction: Medication adherence is a notable problem for patients diagnosed with schizophrenia. A number of psychosocial adherence programs exist to address this problem, yet many require in-person assistance and are typically resource intensive. Drawing on a diverse body of empirical literature, a psychosocial treatment program was designed to help patients turn medication-taking into a habit. This program leverages a call-center with an associated technology interaction portal, text messaging and a "habit kit" that is sent to patients.
In intervention studies, insufficient attention is often given to how programs are 'metabolized' by patients and HCPs, e.g., the personal experience of the intervention that may affect the outcome of the program. This information is important to obtain in order to ensure the goals of the intervention are achieved.
To this end, a treatment intervention refinement study was conducted. The overarching study goals for this pilot were twofold: (1) Gain insight on the ease of implementation, the feasibility of implementation, and the preliminary data on the effectiveness of the program, and (2) Determine modifications for improvement of the program to aid medication adherence habit formation. This poster describes the innovative study methodology that was employed to elucidate how this psychosocial treatment was "metabolized" by patients and call center agents.

Methods: This study included male and female patients with a current diagnosis of schizophrenia or schizoaffective disorder as confirmed by the Principal Investigator (PI). The patients were currently prescribed medication to treat their schizophrenia or schizoaffective disorder, and were 18 to 65 years of age. The study included a screening period and a treatment refinement period (Days 1 to 24). During the treatment refinement period, patients were sent their "habit kit" and participated in six separate phone calls with call center agents. Clinical academic as well as qualitative user-centered design researchers were involved in data collection. Data collection methods included: (i) in-home patient observation when patients were observed during one phone call and first-time interaction with the "habit kit", (ii) interviews which occurred separately with patients and call center agents, (iii) call center audio recording analyses to evaluate usability and fidelity, (iv) engagement metrics regarding call center contact and text messaging, e.g., duration of calls, number of text messages and (v) clinical screening and psychometric measures. Patient interviews and observation took place exclusively in patient homes.

Results: This study is ongoing.

Conclusions: The methods employed in the current study represent a novel approach to the pilot phase of psychosocial treatment program development. These innovative and diverse data collection methods may offer new insights about how psychosocial treatments are metabolized in patients diagnosed with schizophrenia and enable optimal treatment program refinement.

Disclosures: This work was funded by Ostuka Pharmaceutical Development and Commercialization (OPDC). AH, DP, JPD and EL are employees for OPDC. DV is an employee of UT. AA is an employee of GfK.

2 Validation of the Tablet-based Brief Assessment of Cognition (BAC App) for Schizophrenia

Atkins AS[1], Keefe RSE[1, 2], Davis VG[1], Tseng T[1], Vaughan A[1], Harvey P[3], Narasimhan M[4], Patterson T[5]

[1]NeuroCog Trials, Durham, NC, [2]Duke University Medical Center, Durham, NC, [3]University of Miami Miller School of Medicine, Miami, FL, [4]University of South Carolina School of Medicine Columbia, SC, [5]University of California, San Diego, School of Medicine, San Diego, CA

Methodological Question Being Addressed: Determining the validity and usability of a tablet-based adaptation of a cognitive test battery for use in schizophrenia.

Introduction: Computerized tests benefit from automated scoring procedures and standardized administrator instructions. These methods can reduce the potential for rater error. However, especially in patients with severe mental illnesses and neurologic disorders, the equivalency of traditional and tablet-based tests cannot be assumed. The Brief Assessment of Cognition in Schizophrenia (BACS) is a pen-and-paper cognitive assessment tool that has been used in hundreds of research studies and clinical trials, and has normative data available for generating age- and gender-corrected standardized scores. A tablet-based version of the BACS called the BAC App has been developed. This study compared performance on the BACS and the BAC App in patients with schizophrenia and healthy controls. Test equivalency was assessed, and the applicability of paper-based normative data was evaluated.

Methods: Participants included 48 patients (23 female) with schizophrenia and 50 healthy controls (25 female) recruited from three academic sites including the University of California-San Diego, the University of Miami - Miller School of Medicine, and the University of South Carolina. All participants were assessed with the standard pen-and-paper BACS and the BAC App.

Results: In both groups, the distributions of standardized composite scores for the tablet-based BAC App and the pen-and-paper BACS were indistinguishable, and the between-methods mean differences were not statistically significant. The discrimination between patients and controls was similarly robust with the BAC App (d=1.34) and the BACS (d=1.24). The between-methods correlations for individual measures in patients were r>0.70 except Token Motor (r=0.43) and Tower of London (r=0.61). In patients, performance between the test methods was not significantly different on any test except the Token Motor Test. When data from the Token Motor Test were removed, the between-methods correlation of composite scores improved to r=.88 (df=48; P<.001) in healthy controls and r=.89 (df=46; P<.001) in patients, consistent with the test-retest reliability of each measure.

Conclusions: The tablet-based BAC App generates results consistent with the traditional pen-and-paper BACS. These data support the notion that the BAC App can now be used in clinical trials and clinical practice.

Disclosures: R.S.E. Keefe currently or in the past 3 years has received investigator-initiated research funding support from the Department of Veteran's Affair, Feinstein Institute for Medical Research, National Institute of Mental Health, 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, Asubio, Avanir, AviNeuro/ChemRar, BiolineRx, Biogen Idec, BiolineRx, Biomarin, Boehringer-Ingelheim, EnVivo/FORUM, GW Pharmaceuticals, Janssen, Johnson & Johnson, Lundbeck, Merck, Minerva Neurosciences, Inc., Mitsubishi, Neuralstem, Neuronix, Novartis, NY State Office of Mental Health, Otsuka, Pfizer, Reviva, Roche, Sanofi/Aventis, Shire, Sunovion, Takeda, Targacept, and the University of Texas South West Medical Center. Dr. Keefe receives royalties from the BACS testing battery, the MATRICS Battery (BACS Symbol Coding) and the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). He is also a shareholder in NeuroCog Trials and Sengenix. A.S. Atkins currently or in the past 3 years has received funding from the National Institute of Mental Health and is a full time employee of NeuroCog Trials. V.G. Davis, T. Tseng and A. Vaughan are employees of NeuroCog Trials. P. Harvey has received consulting fees from Acadia, Boehringer-Ingelheim, Forum Pharma, Lundbeck, Otsuka America, Sanofi Pharma, Sunovion Pharma, and Takeda Pharma during the past year. He also received a research grant from Takeda. He also receives royalties from the MCCB and for the Brief Assessment of Cognition. M. Narasimhan currently or in the past 3 years has received investigator-initiated research funding support from NIMH, NIDA, NIAAA, NIH, Eli Lilly, Janssen Pharmaceuticals, Forest Labs, and Pfizer. She currently or in the past 3 years has received honoraria, served as a consultant, or advisory board member for Eli Lilly. T. Patterson currently or in the past 3 years has consulted with NeuroCog Trials, Abbott Labs, and Amgen, and has been funded by NIMH.

3 Adjunctive Therapy for Bipolar Depression: Effect on Response of Prospective vs. Retrospective Treatment with Lithium or Valproate Prior to Randomization

Tsai J, Hsu J, Pikalov A, Loebel A

Sunovion Pharmaceuticals, Inc., Fort Lee, NJ and Marlborough, MA, USA

Methodological Question Being Addressed: In a bipolar depression study evaluating antipsychotic treatment adjunctive with lithium or valproate, is there a difference in assay sensitivity depending on whether insufficient response to lithium or valproate is ascertained at initial screening, or demonstrated prospectively during a run-in period prior?

Introduction: Lurasidone has demonstrated efficacy for the treatment of bipolar depression in pivotal studies, both as a monotherapy and as adjunctive therapy with lithium or valproate. A second adjunctive therapy study permitted recruitment of both prospectively treated patients and those already receiving lithium or valproate. We now examine the impact of recruitment strategy on the ability to detect a drug effect.

Methods: Patients with bipolar I depression were randomized to 6 weeks of double-blind treatment with lurasidone (N=180) or placebo (N=176) added to background treatment with lithium or valproate. Prior to randomization, all patients received ≥4 weeks of treatment with lithium or valproate at therapeutic doses utilizing one of the following methods: (1) by initiating prospective treatment after study enrolment (run-in cohort); or (2) by recruitment of patients confirmed to be on therapeutic doses at screening (non-run-in cohort). The primary *a priori* endpoint was change on the Montgomery-Åsberg Depression Rating Scale (MADRS). In a pre-planned analysis, the effect of recruitment strategy (run-in vs. non-run-in) on endpoint treatment response was evaluated.

Results: A comparison of baseline characteristics for the run-in cohort (61.5% of total sample) vs. non-run-in cohort (38.5%) found minimal differences in demographic and clinical variables, with similar MADRS total scores (29.0 vs 29.2). However, among patients in the run-in cohort, those who were randomized showed significantly less screen-to-baseline improvement in the MADRS total score compared with patients who were not randomized (-1.1 vs. -8.6; P<0.001). At week 6 endpoint, improvement in the placebo-subtracted MADRS total score was significantly larger for the non-run-in cohort compared to the run-in cohort (LS mean difference in endpoint change scores, 4.6; P=0.009). Effect sizes were consistently larger for the non-run-in vs. run-in cohort cohorts at weeks 1 (0.28 vs. 0.02), 2 (0.34 vs. 0.12), 3 (0.56 vs. 0.02), 4 (0.52 vs. 0.11), 5 (0.52 vs. 0.07), and 6 (0.31 vs. 0.03).

Conclusions: Endpoint improvement in MADRS total scores for adjunctive lurasidone vs. placebo was significantly larger for patients adequately treated with lithium or valproate prior to screening (non-run-in cohort) compared to patients treated prospectively with lithium or valproate prior to randomization (run-in cohort). This result is consistent with previous reports (Nelson JC, et al. Am J Psych 2009;166:980-91; Iovieno N, et al. J Clin Psych 2012;73:676-83) suggesting that use of a prospective run-in period may reduce the ability to detect a treatment effect as significant in adjunctive (combination therapy) trials in unipolar depression. This result further highlights the methodologic challenge of conducting adjunctive therapy studies in bipolar depressed populations.

Disclosures: Authors are employees of Sunovion Pharmaceuticals, Inc.
ClinicalTrials.gov: NCT01284517 Sponsored by Sunovion Pharmaceuticals, Inc.

4 Examining Effects of Language and Region on Test-Retest Reliability of the MATRICS Consensus Cognitive test Battery (MCCB)

Georgiades A[1, 2], Davis VG[1]; Atkins AS[1], Stroescu I[1], Keefe RSE[1, 2]

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

Methodological Question Being Addressed: The MATRICS Consensus Cognitive Battery (MCCB) is the primary cognitive test battery recommended for clinical trials in schizophrenia populations. It measures several dimensions of cognitive performance and has been validated for test–retest reliability, practicality for clinical trials and tolerability for patients (Nuechterlein et al., 2008). Studies also suggest that the MCCB demonstrates small practice effects (Keefe et al., 2010). The MCCB has been successfully translated to multiple foreign languages and used in many international clinical trials which has generated an interest to test whether language and regional differences have an influence on the psychometric characteristics of the MCCB.

Introduction: The primary aim of the present study was to examine influence of testing language: "English speaking" (98% North American) vs "non-English speaking" (South American, Asian and European) populations, on test-retest stability and practice effects from repeated administration of the MCCB in a large cohort of schizophrenia patients.

Methods: We examined a pooled cohort of 2662 schizophrenia patients from 18 clinical trials. All participants completed the MCCB twice before starting treatment: once at a screening and once as part of the baseline assessment. There were 1815 participants from English speaking countries and 847 participants from non-English speaking countries (131 patients

from South America, 255 patients from Asia and 461 patients from Europe). The composite MCCB score and the 2 cognitive domains, 'Working memory' (WM) and 'Speed of Processing' (SOP) were studied.

Results: Composite Score: The composite scores at screening and baseline showed high and similar test-retest correlations in English speaking (ICC=.90) and non-English speaking regions (ICC's=.87-.92). The effect size of the practice effect between screening and baseline was moderate, with a higher practice effect in the non-English speaking populations (Cohen's d's=.50-.53) compared to the English speaking patients (Cohen's d=.34). Composite scores were higher in English speaking compared to Non-English speaking populations both at screening and baseline (see Table 1)
WM domain: The associations between WM scores at screening and baseline showed moderate/high and similar test-retest correlations in both English speaking (ICC=.81) and the non-English speaking regions (ICC's=.78-.87). The practice effects for WM were small both in the English speaking (Cohen's d=.19) and the non-English speaking (Cohen's d's=.16-.29) populations. WM domain scores were higher in European and Asian populations compared to English speaking and South American populations both at screening and baseline (see Table 1).
SOP domain: The associations between SOP scores at screening and baseline showed high test-retest correlations in both the English speaking (ICC=.83) and non-English speaking population (ICC's=.85-.87). The practice effects for SOP were low/moderate both in the non-English speaking population (Cohen's d's=.43-.51) and the English speaking population (Cohen's d=.31). The SOP domain scores were significantly lower in non-English speaking compered to English speaking population both at screening and baseline (see Table 1).

Table 1. Mean±std T-scores for the Overall composite, the WM domain and the SOP domain at screening and baseline for English speaking and non-English speaking populations (South American, Asian and European).

		Screening Mean±std	Baseline Mean±std
Overall Composite	English Speaking	26.3±12.0	28.2±12.3
	South American	17.9±10.6	20.9±11.5
	Asian	23.0±11.5	26.0±12.3
	European	24.9±12.5	27.3±13.3
Working Memory	English Speaking	34.2±11.4	35.6±11.3
Domain T-score	South American	31.4±10.5	33.5±11.0
	Asian	41.5±11.5	42.6±10.9
	European	36.9±12.4	38.6±12.4
Speed of Processing	English Speaking	31.5±11.8	33.7±11.9
Domain T-score	South American	17.3±12.9	20.3±12.3
	Asian	23.9±12.8	27.4±13.0
	European	25.2±12.8	28.3±12.9

Conclusions: Despite significant differences in MCCB scores by language and region at screening and baseline, the practice effects and test-retest stability of the composite score, WM and SOP domains were similar in English speaking and non-English speaking populations.

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

5 The eCOA Negative Symptom Assessment-16 (NSA-16) Instruction Manual –Version 3.0

Alphs L[1], Velligan D[2], Daniel D[3]

[1]Johnson and Johnson, [2]University of Texas Health Science Center, San Antonio, [3]Bracket Global, LLC

Methodological Question Being Addressed: Rationale for creation of an eCOA version of the semi-structured interview and manual to facilitate accurate and reliable use of the NSA-16.

Introduction: The purpose of the Negative Symptom Assessment -16 is to permit the reliable rating of reduction or absence of emotional expression and volitional behaviors commonly associated with the concept of negative symptoms in schizophrenia. (Axelrod BN, Goldman, RS, Alphs, LD, 1993) The NSA-16 Instruction Manual includes a semi-structured interview and detailed instructions to enhance the accuracy and reliability of assessment. In Version 3.0 we have revised the NSA-16 Manual to include additional instructions and a more detailed semi-structured interview. Version 3.0 has been adapted for either paper or electronic capture (eCOA) on a tablet or desk top computer.

Method: Principles for revision of the NSA-16 Instruction Manual and Semi-structured interview included enhanced clarity of sources of information, reference population, time frame and factors considered in rating the global score. Additional details and probes in the structured interview enhance the consistency and thoroughness of assessment.
Principles guiding the creation of an eCOA version included rapid acquisition of data, lack of need for transcription from paper to eCRF, capacity for real time edit checks for logical inconsistencies in scoring prior to data submission and the capacity for alerts for disqualifying scores after data submission at screening and baseline.

Discussion: The revisions in Version 3.0 of the NSA-16 Instruction Manual are expected to provide further enhancement

of thoroughness of interview technique and accuracy and reliability of ratings. The eCOA version permits paperless data acquisition with incorporation of ratings quality edit checks prior to data submission and disqualification alerts at screening and baseline after data submission. The eCOA version will be piloted in an upcoming global clinical trial.

6 Mismatch Negativity and P300 as Biomarkers of Target Engagement for Transcranial Direct Current Stimulation in Schizophrenia

Dunn W[2], Rassovsky Y[1,2], Wynn J[2], Wu A[3], Iacoboni M[2,4], Chung A[3], Hellemann G[2], Green MF[2,5]

[1]Department of Psychology and Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, Israel, [2]Department of Psychiatry and Biobehavioral Sciences, UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA, [3]Department of Neurology, University of California, Los Angeles, USA, [4]Ahmanson-Lovelace Brain Mapping Center, University of California, Los Angeles, USA, [5]Department of Veteran Affairs VISN-22 Mental Illness Research Education Clinical Center, Los Angeles, CA, USA

Methodological Question Being Addressed: The aim of this study is to determine if EEG event related potential (ERP) measures of Mismatch Negativity (MMN) and P300 can serve as biomarkers of target engagement for transcranial direct current stimulation (tDCS) in schizophrenia.

Introduction: Schizophrenia is an illness characterized by deficits in early sensory and neurocognitive processing. Existing antipsychotic treatments manage positive symptoms such as delusions and hallucinations, but have little impact on sensory processing deficits and neurocognition. Both of these are significant predictors of patient functional outcomes. Transcranial direct current stimulation (tDCS) is a novel non-invasive neuromodulatory technique that has been shown to reduce hallucinations and improve working memory in schizophrenia. In the current study we examined the effects of tDCS on neural measures of basic auditory processing (mismatch negativity) and cognitive processing, (the P300 oddball response) as these measures may serve as biomarkers of tDCS engagement of targeted cortical networks.

Methods: 30 outpatients with schizophrenia were randomized into a single-blinded between subjects study with 10 subjects each receiving bilateral anodal, cathodal, or sham tDCS (active tDCS- 30 min, 2mA). Active stimulation was delivered through two 5X7 cm electrodes (0.02mA/cm2) placed bi-frontally at Fp1 and Fp2 positions (same polarity), with a reference electrode placed on the upper arm. Subjects underwent the ERP protocol at baseline and one week later approximately 1 hr after receiving a single tDCS stimulation. MMN and P300 amplitudes were measured as the mean of activity in the 135-205ms and 250-350ms latency range, respectively.

Results: Anodal stimulation yielded a significant decrease ($p<0.02$) in the MMN amplitude (post-stimulation verses baseline). No statistically significant changes in MMN amplitude were observed for cathodal or sham conditions. The P300 measure did not demonstrate significant changes for any of the three stimulation conditions. All subjects tolerated the stimulation procedure with no reports of adverse events or dropouts.

Conclusions: The change in MMN amplitude after bifrontal anodal stimulation is the first demonstration that a single bilateral session is sufficient to affect change in an auditory ERP signal in schizophrenia. These finding support MMN as a potential biomarker for tDCS engagement of the auditory processing network. Such basic auditory processing contributes to higher order social cognition which has been shown to be correlated to patient functional outcomes. In schizophrenia, there are known deficits in both basic auditory processing and social cognition, and thus treating these deficits could improve patient outcomes. A potential application of tDCS would be to correct the defects in auditory processing, the effects of which would cascade and improve higher order social cognition. The study findings suggest that tDCS stimulation changes the amplitude of MMN response. Known nodes of this network include the pre-frontal cortex which was the target of tDCS in this study. Therefore, in conjunction with the MMN results, there is evidence that MMN could serve as a target engagement biomarker for tDCS treatment of auditory processing deficits in schizophrenia.

7 What Blinded Raters Don't See: Relapse Events in a Bipolar Study Detected By Site-Based Raters Versus a Computer Simulated Rater

Sachs G[1], Peters A[2]

[1]Bracket, [2]University of Illinois at Chicago

Methodological Question Being Addressed: Relapse prevention studies can fail due to lower than expected event rates. We explore alternative explanations for low relapse rates in a recent relapse prevention study for bipolar disorder.

Introduction: Relapse prevention studies are powered based on an estimate of expected events during an interval of double blind follow-up, but lower than expected relapse rate have been observed in several recent bipolar maintenance studies. Some hypothesize this problem simply reflects enrolment of subjects unlikely to relapse. Alternatively, low rates

of relapse events may reflect subtle misalignment of incentives between site raters and subjects motivated to complete the study and the research objective of detecting relapse events. In this case, subjects meeting numerical criteria for "relapse" may seem well enough to deflate scores and enable subjects to remain in study.
To explore these competing hypotheses, we examined two categories of relapse events reported by site-based raters (SBR) and a computer simulated rate (CSR)r: Rating scale defined relapse events (RSDRE) and clinical judgment relapse events (CJRE).

Methods: These analysis used a de-identified database from a global clinical trial evaluating an investigational drug for prevention of relapse, which was stopped for futility. The trial randomized bipolar I subjects in remission at screening and baseline to placebo or one of three doses of the study drug. Tandem ratings were obtained on the Montgomery Asberg Depression Rating Scale (MADRS) and the Young Mania Rating Scale (YMRS) by SBRs and an independent interview administered by a CSR. The protocol defined a relapse event as occurring if the site clinician deemed the subject was in need of a non-study intervention or when either the MADRS or YMRS score was ≥ 16. For purpose of this analysis, the RSDRE were determined for SBR and CSR using the protocol specified scores. For CJRE, the criteria were operationalized using the STEP-BD definition of "Roughening" (≥2 clinically significant symptoms). Symptoms counted as clinically significant when the rater scored a scale item above the mid-point in the range for that item.

Results: The Study randomized 403 subjects in a double-blind placebo controlled study with 12 months of follow-up.
Table1: Relapse events for Site-based Raters and Computer simulated Rater

	N	Rating Scale Defined Relapse Events (%)	P (Chi-Square)	Clinical Judgment defined Relapse Events (%)	P (Chi-Square)
SBR*	403	82 (20.3%)	<.01 (35.93)	64 (15.9%)	<.01 (20.70)
CSR	403	160 (39.7%)		118 (29.3%)	
SBR or CSR	403	167 (41.4%)	<.01 (41.99)		

* The occurrence of any relapse event detected by SBRs resulted in termination. Only 7 SBR detected events occurred without a preceding or concurrent event detected by the CSR.

Conclusions: For both RSDREs and CJDREs, we found the CSR detected relapse events nearly twice as frequently at SBRs. Our results favour the hypothesis that subtle bias in reporting or rating reduced the number of events detected by SBR. Further studies are needed to evaluate the potential of computer simulated raters.

Disclosures: Gary Sachs is a full Time employee of Bracket. Amy Peters has no disclosures.

8 Elevated Translocator Protein (TSPO) in Subjects at High Risk of Psychosis and in Schizophrenia: an [11C] PBR28 PET Brain Imaging Study

*Selvaraj S[1,2,8], *Bloomfield PS[1,2], Veronese M[3], Rizzo G[4], Bertoldo A[4], Owen DR[5], Bloomfield MAP[1,2], Bonoldi I[6], Kalk N[5], Turkheimer F[3], McGuire P[6], de Paola V[7], Howes OD[1,2, 6]

[1]Institute of Clinical Science, Imperial College London, [2]Department of Psychiatric Imaging, Imperial College London, [3]Centre for Neuroimaging Sciences, King's College London, ·University of Padova, Padova, Italy; [5]Imperial College London, [6]King's College London, Institute of Psychiatry, Psychosis Studies King's College London, [7]Department of Neuroplasticity and Disease, Imperial College London, [8]University of Texas Health Science Center at Houston, Houston, Texas

Methodological Question Being Addressed: Efficacy of anti-inflammatory agents in the treatment of psychotic and mood disorders is promising. The underlying pathophysiological mechanisms and specific predictors of anti-inflammatory treatment response are not known. Using [11C] PBR28 and positron emission tomography (PET) imaging, we investigated whether the individuals with ultra high-risk (UHR) symptoms for psychosis have elevated brain microglial activity or not.

Introduction: Abnormal brain immune responses have been implicated in the pathophysiology of schizophrenia and proposed as a mechanism associated with the brain volume decreases and illness-progression seen in schizophrenia. Microglial cells are the resident immune cells of the central nervous system. Microglia activation can be measured in vivo with positron emission tomography (PET) using radioligands specific for the 18KD translocator protein (TSPO),

expressed on microglia associated with inflammation (1). Recent PET in-vivo brain imaging studies show elevated TSPO binding in patients with schizophrenia (2, 3).
However it remains unclear how this relates to the onset of psychotic illness. To determine whether total grey matter TSPO is altered in individuals with ultra high-risk (UHR) symptoms for psychosis. We also seek to determine how this compares with diagnosed patients with schizophrenia.

Methods: We recruited fourteen subjects with UHR symptoms and 14 patients with schizophrenia. Two groups of age, and TSPO genotype matched control subjects were recruited from the same community. All study participants underwent a PET scan with [11C] PBR28, a TSPO ligand and a high resolution MRI scan. The main outcome measure was total grey matter [11C] PBR28 distribution volume ratios (DVRs). DVR is the ratio of the Volume of distribution VT in the regions of interest to VT in the whole brain.

Results: Multiple analysis of variance demonstrated an elevation in total grey matter [11C] PBR28 DVRs in both UHR subjects (controls 2.03 (0.02), UHR 2.09 (0.02); $F=10.33$; $p=0.004$) and patients with schizophrenia (controls 2.47 (0.02), patients 2.56 (0.01); $F=20.8$; $p<0.001$). UHR symptoms, as measured with the Comprehensive assessment of the at risk mental state (CAARMS) were positively correlated with total grey matter [11C] PBR28 DVR ($p<0.01$) in UHR subjects.

Conclusions: Increases in [11C] PBR28 in UHR subjects and patients with schizophrenia suggest that neuroinflammatory changes are associated with the development of psychosis. To our knowledge, these data represent the first evidence of elevated brain microglial activity in people at ultra high risk of psychosis, and show that greater microglial activity is associated with greater symptom severity. The finding raises the possibility that anti-inflammatory treatment may be effective in preventing the onset of the disorder and [11C] PBR28 as a potential biomarker for brain inflammatory-immune response.

Disclosures: Supported by a Medical Research Council (United Kingdom) grant to Dr. Howes (grant number MC-A656-5QD30)and Maudsley Charity grant (grant number 666) to Dr. Howes, as well as the National Institute for Health Research (NIHR) BiomedicalResearch Centre at South London and Maudsley NHS Foundation Trust and King's College London. Mr. Bloomfield has conducted research funded by the Medical Research Council (United Kingdom), the National Institute of Health Research (United Kingdom), and the British Medical Association. Dr. Kalk received funding for a Ph.D. degree by a Wellcome Trust GlaxoSmithKline fellowship, and she has received grant support from GlaxoSmithKline to attend educational events and conferences. Dr. Howes has received investigator-initiated research funding fromand/or participated in advisory/speaker meetings organized by AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Janssen, Lundbeck, Lyden-Delta, Otsuka, Servier, and Roche. All other authors report no financial relationships with commercial interests.

References: 1. Karlstetter M, Nothdurfter C, Aslanidis A, Moeller K, Horn F, Scholz R, et al. (2014): Translocator protein (18kDa) (TSPO) is expressed in reactive retinal microglia and modulates microglial inflammation and phagocytosis. Journal of Neuroinflammation. 11:3.
2. Van Berckel BN, Bossong MG, Boellaard R, Kloet R, Schuitemaker A, Caspers E, Luurtsema G, Windhorst AD, Cahn W, Lammertsma AA, Kahn RS. Microglia Activation in Recent-Onset Schizophrenia: A Quantitative (R)-[11C]PK11195 Positron Emission Tomography Study. Biological Psychiatry. 2008;64:820- 822.
3. Doorduin J, de Vries EFJ, Willemsen ATM, de Groot JC, Dierckx RA, Klein HC. Neuroinflammation in Schizophrenia-Related Psychosis: A PET Study. Journal of Nuclear Medicine. 2009;50:1801-1807.
The above data is published as *Peter S. Bloomfield, *Sudhakar Selvaraj, Mattia Veronese, Gaia Rizzo, Alessandra Bertoldo, David R. Owen, Michael A.P. Bloomfield, Ilaria Bonoldi, Nicola Kalk, Federico Turkheimer, Philip McGuire, Vincenzo de Paola, Oliver D. Howes. Imaging translocator protein (TSPO) in subjects at high risk of psychosis and in schizophrenia: an [11C] PBR28 PET brain imaging study. Am J Psychiatry. 2015 Oct 16:appiajp201514101358. [Epub ahead of print] Equal contribution. Joint 1st Authors.

9 Early Indicators of Poor Data Quality in Schizophrenia Clinical Trials

Kott A, Daniel D

Bracket

Methodological Question Being Addressed: We intend to examine whether the presence of data quality issues after randomization is associated with the presence of data quality issues identified during the screening period.
Introduction: A primary focus of risk-based data quality monitoring program in clinical trials is early identification and remediation of issues that may detract from signal detection. Ideally such a program should identify problematic ratings before a subject is randomized into the study. In the current analysis, we examined whether the incidence of data quality issues after randomization was affected by the presence of data quality issues in the screening period.

Method: Using negative binomial regression we analyzed blinded data from 14 international double blind placebo controlled schizophrenia trials involving 10,056 subjects (67,584 visits). The incidence of data quality issues after randomization (logical inconsistencies in scoring among PANSS items; unusually large changes in the PANSS, erratic changes in the PANSS and identical scoring of 30/30 PANSS items across consecutive visits; inconsistencies between CGI-I scores and changes in either the PANSS or CGI-S from baseline) was compared between those subjects who recorded a data quality issue in the screening period and those who did not.

Results: The incidence of individual type post-randomization data quality issues was significantly affected by the presence of the same type data quality issues in the screening period. For example the presence of logical inconsistencies in scoring among PANSS items at screening increased the expected incidence rate of these within PANSS inconsistencies almost 8 times. Similarly, the incidence rate of identical scorings on the PANSS in the post-baseline data increased almost 7 times if identical ratings were present at baseline. The presence of some quality indicators in the screening period had a significant effect on the incidence of other type of data quality issues in the post-randomization period. For example, the presence of large changes in PANSS from screening at baseline increased the incidence rate of erratic ratings in the post-randomization period almost 4 times. The presence of any data quality issues in the screening phase increased the incidence of any data quality issues after randomization more than 2 times.

Conclusions: Our data show strong and significant effect of the presence of data quality issues in the screening phase on the incidence of the same or related data quality issues after randomization. This represents an important finding because it indicates that raters who commit serious rating and interviewing errors in the screening phase are likely to commit significantly more of the same or related errors later in the study with the potentially negative effect on signal detection. The findings as well stress the necessity to intervene early to prevent the identified raters to repeat the errors.

Disclosure: One or more authors report potential conflicts. Both authors are full time employees of Bracket, LLC.

10 How do Key Co-Primary Measures of Functional Capacity Predict Real World Function in Schizophrenia?

Keefe RSE[1, 2], Atkins AS[1], Davis VG[1], Harvey P[3,] Narasimhan M[4], Patterson T[5]

[1]NeuroCog Trials, Durham, NC, [2]Duke University Medical Center, Durham, NC, [3]University of Miami Miller School of Medicine, Miami, FL, [4]University of South Carolina School of Medicine, Columbia, SC, [5]University of California, San Diego, School of Medicine, San Diego, CA

Methodological Question Being Addressed: Determining the relative ability of measures of functional capacity that are used as co-primary measures in schizophrenia cognition clinical trials to predict real world functioning.

Introduction: Patients with schizophrenia have profound and disabling cognitive deficits that interfere with multiple aspects of daily functioning. Indeed, research suggests that cognitive impairment accounts for more disability in real world functioning than any other aspect of the illness, including psychosis (e.g. August, 2012). In order to assess the potential impact of cognitive enhancement therapy on functioning, the FDA has required that clinical trials for cognitive impairment in schizophrenia demonstrate improvement on a standard performance-based cognitive assessment, as well a co-primary measure of 'functional capacity' that can serve as an intermediary between cognitive and functional improvement, and may signal increased potential for improved outcomes. Although the relationship between cognitive performance and standard measures of functional capacity have been well-described, less attention has been devoted to the relationship between functional capacity and measures of real world function. In order explore this, we examined the relationship between real world function and performance on three measures of functional capacity, including the University of California San Diego Performance-Based Skills Assessment (UPSA-VIM), the Schizophrenia Cognition Rating Scale (SCoRS) and Virtual Reality Functional Capacity Assessment Tool (VRFCAT).

Methods: Participants included 158 patients who met DSM-IV TR criteria for schizophrenia. Participants were recruited as part of a non-treatment psychometric validation study conducted across three research sites, including the University of South Carolina, the University of Miami - Miller School of Medicine, and the University of California, San Diego.
All subjects completed the MATRICS Consensus Cognitive Battery (MCCB), UPSA-VIM, SCoRS and VRFCAT at the same study visit. Real world function was evaluated using the Specific Levels of Functioning (SLOF; Schneider & Struening, 1983). The same informant was used for both the SCoRS and SLOF measures. The SLOF dependent variable for the statistical analyses was the total score across the all subscales. Correlations between the SLOF, MCCB composite score and each measure of functional capacity were assessed using Pearson correlation coefficients.

Results: All three measures of functional capacity demonstrated significant correlations with SLOF total score. Consistent with prior findings, the correlation between cognition (MCCB) and real world function as measured by the SLOF was

modest but significant, r=.33, p<.001. Of the three functional capacity measures, the SCoRS demonstrated the strongest correlations with the SLOF, r=-.57 (p<.001), a finding likely influenced by informant input to both measures. The SCoRS correlation with the MCCB was r=-.42 (p<.001). UPSA-VIM correlations with the SLOF and MCCB were r=.27 (p=.002) and r=.70 (p<.001), respectively. The VRFCAT also demonstrated significant correlations with the SLOF, r=.23 (p=.007) for total time and r=.35 (p<.001) for total errors. Correlations between the VRFCAT and MCCB composite were r=.57 and r=.39 for total time and total errors, respectively (p<.001 for both).

Conclusions: Current measures of functional capacity in schizophrenia demonstrate modest-to-moderate correlations with assessment of real world function. Performance-based functional capacity measures have stronger correlations with performance based measures of cognition, while the interview-based measure of functional capacity had a stronger correlation with a real-world functional scale. Results are consistent with the conception of functional capacity as a potential mediator of the relationship between cognition and real world function.

Disclosures: R.S.E. Keefe currently or in the past 3 years has received investigator-initiated research funding support from the Department of Veteran's Affair, Feinstein Institute for Medical Research, National Institute of Mental Health, 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, Asubio, Avanir, AviNeuro/ChemRar, BiolineRx, Biogen Idec, BiolineRx, Biomarin, Boehringer-Ingelheim, EnVivo/FORUM, GW Pharmaceuticals, Janssen, Johnson & Johnson, Lundbeck, Merck, Minerva Neurosciences, Inc., Mitsubishi, Neuralstem, Neuronix, Novartis, NY State Office of Mental Health, Otsuka, Pfizer, Reviva, Roche, Sanofi/Aventis, Shire, Sunovion, Takeda, Targacept, and the University of Texas South West Medical Center. Dr. Keefe receives royalties from the BACS testing battery, the MATRICS Battery (BACS Symbol Coding) and the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). He is also a shareholder in NeuroCog Trials and Sengenix. A.S. Atkins currently or in the past 3 years has received funding from the National Institute of Mental Health and is a full time employee of NeuroCog Trials. V.G. Davis is a part-time employee of NeuroCog Trials. P. Harvey has received consulting fees from Acadia, Boehringer-Ingelheim, Forum Pharma, Lundbeck, Otsuka America, Sanofi Pharma, Sunovion Pharma, and Takeda Pharma during the past year. He also received a research grant from Takeda. He also receives royalties from the MCCB and for the Brief Assessment of Cognition. M. Narasimhan currently or in the past 3 years has received investigator-initiated research funding support from NIMH, NIDA, NIAAA, NIH, Eli Lilly, Janssen Pharmaceuticals, Forest Labs, and Pfizer. She currently or in the past 3 years has received honoraria, served as a consultant, or advisory board member for Eli Lilly. T. Patterson currently or in the past 3 years has consulted with NeuroCog Trials, Abbott Labs, and Amgen, and has been funded by NIMH.

11 Assessing Change in Neurocognition: Reliable Change Indices for Neurocognitive Assessment in Schizophrenia

Khan A[1,3], Lindenmayer JP[1,3], Lucic L[4], Ozog V[3,5], Fregenti S[3,5], Goldring A[3], Thanju A[3]

[1]Nathan S. Kline Institute for Psychiatric Research, [2]New York University School of Medicine, [3] Manhattan Psychiatric Center, [4]Pratt Institute, [5]Columbia University

Methodological Question Being Addressed: Currently, there are varied definitions of improvement in neurocognition in patients with schizophrenia or related disorders. What is the most optimal measure to assess change and improvement for neurocognitive data? The study will provide a comparison of concepts and analysis of clinical significance and the reliable change index (RCI) using pre and post MCCB-MATRICS scores, in order to determine the amount of test score change that is necessary to be deemed statistically reliable.

Background: A particularly important role of neurocognitive assessment is measuring change in cognitive functioning over time. The MCCB-MATRICS is the widely used to assess neurocognition in schizophrenia. Clinical efficacy of an intervention relative to a placebo or control condition is generally confirmed through statistically significant differences. However, statistical significance does not in itself provide concise information about an intervention's clinically meaningful effects. Methods for measuring change have been discussed and continue to appear in the literature, with no optimal conclusion, thus researchers are using varying definitions and estimates for change scores. Examples of change score methods include the standard deviation (S.D.) method, reliable change indices (RCI), standardized-regression-based (SRB) methods, percent change, and a specified amount of domain change. As an attempt to develop a standard method of estimating significant change (statistically and clinically), we propose comparing various change score methods employed in neurocognition and comparing them to Reliable Change Index (RCI). The aim of this study is to provide a comparison of concepts and analysis of clinical significance and the reliable change index (RCI) using pre and post

MCCB-MATRICS scores, in order to determine the amount of test score change that is necessary to be deemed statistically reliable.

Methods: Data from a parent study examining neurocognitive change following cognitive remediation therapy in patients with schizophrenia were examined. Three methods of RCI (Jacobson-Truax, Edwards-Nunnally, and Hageman-Arindell methods) were compared to published methods of neurocognitive change, SD change, effect size change, percent domain change, and SRB methods.

Results: For the three RCI methods, 40.56%, 41.23% and 41.33% showed reliable improvements in at least one domain of the MCCB-MATRICS. For SD change, 76.97% showed improvement, 69.73% improved on at least one domain looking at percent change (published method), and 68.11% showed improvement with effect size change, and 81.23% showed improvement with >20% improvement in at least one domain. When comparing RCIs with previously published change score methods, only 16.53% of improvement also resulted in RCI significant improvement. Regarding clinically meaningful improvement, the Hageman-Arindell method was most concordant with all three RCI measures as this method differentially analyzes clinically meaningful change at the individual level and at the group level (i.e., obtaining proportions of patients who have reliably changed and passed the cutoff point).

Conclusions: Our findings that all the strength of association of varied change score by measure and method can significantly affect efficacy outcomes, and reporting both RCI and clinical meaningful change is warranted. Outcome studies often assess statistically significant change, which may not be clinically meaningful. Since the pattern of cognitive change occurring early in the course of schizophrenia may be particularly useful, future research may examine cognition score change in relation to change observed on CT or MRI scans may yield particularly useful prognostic and treatment information about individuals with schizophrenia.

Disclosures: Authors do not have any disclosures to report.

12 Assessing Clinician's Subjective Experience with Psychometric Tools for Suicide Assessment

Dallabrida S[1], Yavorsky C[2]

[1]eResearch Technologies, [2]CRONOS CCS

Methodological Question Being Addressed: While many clinical trials, beyond CNS, are now including a suicide ideation and behavior (SIB) scale, little study has b5een devoted to the clinician's subjective experience with the patient and with commonly used SIB measures, which is highly valued by the phenomenological traditions. Additionally, with the development of self-reported, clinician reported and 'dual' reported tools for SIB, the algorithm for the decision making process for which psychometric evaluation to use, whether it is dependent on indication, and the levels of complexity of scales remain unclear.

Introduction (Aims): (1) to assess clinician's views on self-reported, clinician reported, dual reported SIB measures, (2) to assess clinician's experience with currently used measures, (3) to establish models for the decision making process for evaluating suicide risk.

Methods: We used a qualitative research approach, interviewing psychiatrists, psychologists or clinicians in the field of schizophrenia, bipolar, depression/anxiety or related disorders, who have experience using SIB measures. 112 general practice, pharmaceutical, academic and government-based clinicians in the United States participated. Surveys were deployed in English via Survey Monkey™ and consisted of 18 questions assessing: Clinician Demographics, Current SIB measure utilized, views on complexity of SIB measure, views on SIB reporting in Clinician-reported and Self-Reported measures, collateral information, cognitive impairment and reporting, acceptability of self-reported measure per disease indication.

Results: A majority of the clinicians who completed the survey were Psychologists (33.33%; clinicial/hospital practice/research). There were 22.22% Psychiatrists (clinicial/hospital practice/research) and 11.11% psychiatrists from the pharmaceutical and biotechnology industry. More than 50% of the sample had > 10 years' experience. 55.56% use the C-SSRS and 33.33% each used the InterSept and Sheehan-STS scales. 22.23% reported using the electronic C-SSRS. 77.78% of clinicians reported deficiencies in the current scale they are using. Of these, 35.54%, 29.34% indicate the scale is too complex for the STS and InterSept, respectively. The largest percentage of clinicians (43.23%) reported the C-SSRS was "not effective for the population they serve." When comparing clinicians views on self-reported and clinician-reported SIB assessment, 22.23% and 33.33%, respectively, agree that their patients provide accurate reports. Only 33.33% of clinicians believe that patients will report more information on a self-reported measure, however many clinicians commented that patients who are suicidal are less likely to self-report. For oncology related diseases, multiple sclerosis, personality disorders, dermatology and metabolic abnormalities, 100% of clinicians agree a self-reported SIB assessment
is acceptable.

Conclusion: Self-reported questionnaires may reveal higher frequency and severity of SIB than clinician-reported in specific disease areas. Results suggest that for most disease areas, clinicians suggest that a combination of self-reported and clinical-reported SIB assessment would be beneficial. Although additional investigation is needed, especially from non-CNS clinicians, results confirm that currently existing measures and algorithms for SIB can benefit from refinement.

13 Advancing Trial Design Methodology to Evaluate Drug Treatment Effects on Suicidal Ideation and Behavior

Di Cesare F[1], Stewart M[2] on behalf of the ISCTM Working Group on Trial Design and Methodology for SIB Treatment

[1] Leoben Research ltd, UK, [2] Pfizer, USA

Methodological Question Being Addressed: Trial design methodology to evaluate drug treatment effects on suicidal ideation and behavior.

Introduction: As the effective prevention and treatment of suicidal ideation and behavior (SIB) still remains a largely unmet medical need, clinical drug research for SIB treatment has been characterised by overall low activity (< of 30 new trials designed and completed worldwide in the past 15 years – source: International Clinical Trial Registry Platform of the World Health Organisation, July 2015), mostly sponsored/funded by public institutions, high heterogeneity in trial design, limited sample sizes, ample variability in the selection criteria and clinical characterisation of study populations, as well as broad diversity in clinical assessment strategies and methodologies to evaluate treatment outcome(s). Current limitations in trial design and methodologies may represent a barrier to the development of effective drug treatments for SIB.

Method: Within the scope of activities of the ISCTM Working Group on Suicidal Ideation and Behaviour Assessment, an "ad hoc" international group of 21 subject-matter experts from Academic, Pharmaceutical, and Clinical Research institutions met on a regular basis to address specific topics from June through November 2015. Objectives of the WG included the review of current methodological approaches and the development of consensus on the key medical and scientific requirements for design of clinical studies and programs to be used for regulatory approval of therapies to treat SIB. In particular, priority was given to issues preventing the advancement of knowledge and technology development in the field. The scope of review included: selection of the study population, demonstration of clinical benefits and evaluation of efficacy, experimental designs for SIB trials, the choice of control and the use of placebo in SIB clinical trials, safety risk assessment and mitigation strategies in SIB trials, biomarkers applied to SIB drug development. The WG initially delivered a list of methodological issues/points for consideration in need of consensus within the broader topic areas; then developed list of priority issues (and related statements) through an iterative group review process; twenty one statements were included in the pre-consensus meeting survey for further validation; finally, upon WG review of survey results and the methodological points considered during the WG preparatory work, the final list of highly prioritised issues and the related statements were developed for open discussion at the ISCTM-SIB consensus meeting which was held in Arlington, Virginia, 17-18 Nov. 2015.

Results: A large consensus was reached on a significant number of issues. The harmonisation of current nomenclatures on SIB would accelerate the advancement of the field. A publication strategy is now in place to disseminate WG findings and conclusions to the global medical, scientific, and regulatory community to promote further debate and methodological advances in the field.

Conclusions: Further debate and continuous interaction between primary stakeholders (medical and scientific, regulatory, SIB treatment developers, patients and families) should be encouraged and actively pursued as an essential strategy for the design and development of optimal treatment strategies for patients.

14 Comparison of Recall Periods and Classification Schemes in Report of Suicidal Ideation and Behavior in Clinical Trials

Dubrava S, Darekar A, Stewart M, Chappell P

Pfizer, Inc.

Methodological Question Being Addressed: To evaluate the impact of two different recall periods (Lifetime and Screening Period) and two versions of the Columbia Classification Algorithm of Suicide Assessment (C-CASA) categories on the report of suicidal ideation and behavior (SIB) events during clinical trials.

Introduction (Aims): FDA requires prospective assessment of SIB in clinical trials for psychiatric and neurologic drugs, and drugs for non-psychiatric indications with potential risk. The Columbia Suicide Severity Rating Scale (C-SSRS) is often used and is recommended by the FDA because C-SSRS data can be directly mapped to the C-CASA event categories. Estimates of the incidence of SIB in clinical trials may be affected by the definition of the pre-study recall period (ie lifetime vs screening period) and by the version of the FDA C-CASA employed. The aims of this study were to investigate the impact of different definitions of the pre-study recall period and the C-CASA classification system on reported SIB in clinical trials.

Methods: Data from 4,610 subjects from 11 studies (4 schizophrenia or bipolar disorder, 3 depression, and 4 GAD studies) were included for a pooled analysis across all treatment groups. Subjects with missing/incomplete data (approximately 6%) were excluded from analysis. Incidence rates and Mantel-Haenszel (M-H) odds ratios were calculated for comparisons between recall periods. Post-Baseline SIB events were modeled using an additive Poisson model adjusted for time on study to determine the number of events per patient-year for each group. Risk differences (95% Cis) of the risk comparison estimates were also determined.

Results: At least one SIB event was reported by 30.3%, 8.1%, and 11.8% of subjects for Lifetime, Screening Period, and Post-Baseline periods, respectively. There were no completed suicides. A positive lifetime SIB history was related to an increased risk of SIB events during the Screening Period. An estimated risk difference (and CI) greater than zero suggested an increased risk of Post-Baseline SIB for subjects who reported any SIB during the Lifetime or Screening Period compared to those with none. The risk of Post-Baseline SIB was greater for those with SIB in the Screening Period. These two recall periods may result in different distributions of baseline SIB severity, with fewer severe events reported during the Screening Period compared to Lifetime. The extended C-CASA 2012 categories identified baseline subgroups with varying levels of risk, but no clear relationship between severity at baseline and higher risk of Post-Baseline events emerged**.**

Conclusions: Post-baseline SI was fairly frequent however SB events were rare in these populations. More recent history (ie, Screening Period) appears to be more strongly linked to an increased risk of SIB during the trial than prior (ie, Lifetime) history. The detailed C-CASA 2012 categories yield baseline subgroups with different risks of Post-Baseline SIB although it was difficult to see a clear pattern between baseline severity and SIB risk. These findings indicate that when evaluating SIB event rates from different studies or programs it is important to assess which baseline recall period and C-CASA categories were used. Results may provide guidance on which patients need closer monitoring during study conduct.

Disclosures: All authors are all full-time employees of Pfizer. The authors gratefully acknowledge the contribution of Ching-Wei Ching during his internship at Pfizer.

15 Options to Address Non-adherence through Subject Involvement

Schueler P[1], Thilaganathan J[2], O'Donohoe P[2]

[1] ICON Clinical Research, [2] CRF Health

Methodological Question Being Addressed: As technology continues to develop, clinical trial methodologies are trying to keep up. However, every change also introduces a potential risk for the study data reliability and its acceptance by regulators, as detailed by EMA [1].
While electronic Clinical Outcome Assessments (eCOA) are frequently used to support endpoints in CNS research, we assessed the options to more actively involve patients in the planning of any study, thus improving patient engagement.

Introduction: CNS diseases are frequently of chronic and progressive nature. Therapies under development in most instances offer a symptomatic relief or a mild disease-modifying effect. These are conditions which traditionally have low adherence to the experimental therapy, what may further dilute any existing therapeutic effect.
eCOA data has become available in real-time. It already is standard to allow site staff immediate access to blinded data to monitor the appropriate use of the e-diaries. As a measure of adherence, compliance rate can get easily calculated (expected assessment vs. actual assessment). However, any corrective feed-back to study subjects is typically given in a non-structured fashion by the site staff, and positive feed-back is usually not provided at all.

Methods: Latest available technology for eCOA documentation was evaluated to see what extent it would allow for a stronger involvement of study subjects in the following areas:

1. Can information be shared with subjects during the course of the study?
2. If so, what data are eligible without introducing bias?
3. What parameters should be measured in future CNS studies to assess the impact on adherence?

We performed a risk analysis of data available in an eCOA device, such as: change in reported level of symptoms, compliance, scheduled events.

Results: We identified several items that can be shared back with the subject regularly without introducing any evaluation bias:

- Subject enrollment status
- Level of adherence
    - For the individual
    - For the full study population
- Number of visits completed versus still outstanding
- Subject activities required until and for next visit

In a pilot project (350 randomized subjects in study in agitated Alzheimer's disease) above measures helped to reach a compliance rate as high as 99% for the 2 key efficacy endpoints.
Not eligible for sharing with the subjects prior to last visit are any data related to safety of efficacy assessments. Providing this level of detail might impact subject's future reports. These items should however be shared at study end.

Conclusions: Technology and regulation allows us to share certain data back with the trial subjects, while other data need to remain confidential during the course of participation of the study subject. The impact of sharing certain data on subjects' willingness to provide most reliable and correct data should be positive, but has to be assessed in a controlled study setting. We plan to introduce this concept of "active involvement" in a coming trial, with half of the subjects being randomized to this concept and the other half randomized to the conventional approach.

References: [1] EMA/INS/GCP/454280/2010; Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials.

16 Mitigating the Effects of Nonadherence in Clinical Trials: Findings of the ISCTM Working Group

Shiovitz TM[1,2], Bain EE[3], Oseas RN[1,2]

[1]California Neuroscience Research Medical Group, Sherman Oaks, CA, [2] CTSdatabase, LLC, Beverly Hills, CA, [3]Global Pharmaceutical Research and Development, AbbVie, Inc., North Chicago, IL

Methodological Question Being Addressed: What is the impact of nonadherence on clinical trials and what recommendations did the International Society for CNS Clinical Trial Methodology (ISCTM) Working Group formulate to identify and mitigate its negative effects?

Introduction (Aims): Nonadherence, the degree to which patients fail to take medications and/or adhere to treatment recommendations as prescribed by their health care providers, is a major public health concern. Nonadherence during the conduct of a clinical trial includes real-world nonadherence as well as *artifactual* nonadherence, subject behaviors unique to the clinical trial setting. The ISCTM Working Group on Mitigating Nonadherence in Clinical Trials explored and modeled nonadherence to put forth specific recommendations designed to address and reduce its negative impact.

Methods: The ISCTM Nonadherence Working Group was established in June, 2014; ISCTM members and NIH scientists were invited by the co-chairs to participate and investigate aspects of nonadherence based on their experience. The Working Group conducted literature searches and had regular meetings where the impact of Duplicate and Professional

subjects and other nonadherence behaviors on clinical trials were presented. Methods to detect and mitigate nonadherence such as biomarkers, subject registries, and medication adherence technology were discussed.

Results: Ultimately, 15 Working Group recommendations were formulated and a paper on the findings was accepted for publication by The Journal of Clinical Pharmacology on November 27, 2015. These recommendations will be presented here.

Conclusion: The 15 recommendations will be presented in their entirety and include: 1) the use of subject registries, PK sampling, biomarkers, and adherence technologies to detect nonadherence, 2) pre-specifying nonadherence criteria for stratification and inclusion in the final analysis, and 3) utilizing adherence data to make decisions on the early term of deceptive subjects and to inform go/no-go decisions in later studies.

Disclosures: Conflict of Interest: Thomas M. Shiovitz is president of CTSdatabase, LLC a clinical trial subject registry discussed in the review.

17 The ADAS-Cog Name Item: More Trouble than it is Worth?

Negash S[1], Weber C[1], Garzio L[1], Randolph, C[1,2]

[1]MedAvante, Inc, [2]Loyola University Medical Center

Methodological Question Being Addressed: Does the 'Name' item in the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-Cog) contribute to the total score, or is it more trouble that it is worth by creating central oversight issues that conflict with protected health information disclosure?

Background: The Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-Cog) is the most widely used primary endpoint in AD trials. The high error rates in administration and scoring of this scale, however, have led to the increased use of centralized oversight, including recording and review of these assessments. This quality control practice is complicated by the 'Name' item on the Orientation subtest, which requires the subject to provide his/her own name. This has implications with respect to release of personal health information, which can occur even in the absence of recording, if the rater mistakenly writes down the name on source documentation.

The Coalition Against Major Diseases (CAMD), a public-private-partnership program at Critical Path (C-Path), has launched C-Path Online Data Repository (CODR) in an effort to accelerate advances in AD therapeutic development. The database contains data from placebo arm of over 20 clinical trials of AD and MCI. The ADAS-Cog data from this database were reviewed to examine the 'Name' item from the Orientation subscale. These data were also augmented with assessments from two double-blind, placebo-controlled AD clinical trials that underwent central review by MedAvante clinicians. Our goal was to examine the error rates on the 'Name' item of the ADAS-Cog, and the extent to which a score on this item contributes to the orientation score and also the total score.

Methods: Four studies in the CODR database generated a total of 6,185 responses to the 'Name' item question, while aggregated data from MedAvante cohort generated 6,669 responses. Assessments were reviewed to determine the percentage of responses to the name question that were scored as 'incorrect'. The correlations between the 'Name' item score and the orientation score (and total score) were examined.

Results: For both cohorts, the percentages of errors on the 'Name' item were very low (less than 1.5%). The item also showed weak correlations with the total score for both CODR ($r = 0.215$) and MedAvante ($r = 0.138$) cohorts. Further, Chronbach's alpha for the 8 orientation items was 0.93, and removing the 'Name' item did not affect the alpha value.

Conclusions: The 'Name' item in the ADAS-Cog has a very low error rate; as a result, it does not contribute to the variance in the total score and also has weak internal consistency with the other orientation items. Further, it creates central oversight issues, conflicting with protected health information disclosure. Thus, it is recommended that the ADAS-Cog instrument be modified to eliminate the name item from the orientation subscale, and simply providing credit for that item in calculation of the total score. This will enable centralized oversight of the administration of this scale without the potential exposure of personal health information.

18 Repurposing Drugs for Agitation and Aggression in Alzheimer's Disease Using a Quantitative Systems Pharmacology Platform

Spiros A[1], Geerts H[1,2]

[1]In Silico Biosciences, Berwyn, PA, [2]Perelman School of Medicine, Univ of Pennsylvania

Methodological Question Being Addressed: Developing high-capacity screen platform for identifying and repurposing potential interesting drugs for treatment of agitation and aggression in Alzheimer's disease. Develop virtual patient platform for designing fast and efficient Proof-of-concept clinical trial.

Introduction: Neuropsychiatric symptoms are important drivers for the cost of treatment of Alzheimer's disease, but there are no good animal models available for drug discovery. Because these symptoms are at the intersection of psychiatry and neurology, repurposing CNS active drugs that act on neurotransmitter pathways makes sense.

Methods: We developed a mechanism-based complex Quantitative Systems Pharmacology computer model of the interaction between ventromedial cortical regions involved in emotion processing and the basal ganglia of the meso-limbic circuit. This advanced model features biophysical realistic representations of neurons firing action potentials that are modulated by G-protein coupled receptors. After implementing AD pathology (gradual loss of synapses and neurons and decreased cholinergic tone) we tested the digital signature of therapeutic interventions that have been studied in agitation and aggression.

Results: We identified a conceptually attractive entropy or information bandwidth measure in specific basal ganglia subregions of the closed cortico-striatal-thalamo-cortical loop that is sensitive to AD pathology and can be restored by different interventions, such as the antipsychotics risperidone, haloperidol and olanzapine, serotonin transport inhibitors such as citalopram, AChE- inhibitors (donepezil and galantamine) and NMDA modulators such as memantine and dextromethorphan.

Discussion: Using a humanized computer model of the Alzheimer's disease brain, we have identified a key neuronal circuit that is related to agitated and aggressive behavior and is sensitive to therapeutic interventions. In principle this system allows one to screen the pharmacological profile of a large number of drug libraries, including drugs for repurposing. The platform can also simulate the impact of comedications and genotypes for supporting clinical trial design.

19 A Human Laboratory Paradigm as a Model to Assess Alcohol Addiction Phenotypes Within a Research Domain Criteria (Rdoc) Framework

Stangl BL[1], Kwako LE[1], Sinha R[2], Ramchandani VA[1]

[1]Section on Human Psychopharmacology, NIAAA/NIH, Bethesda, MD, [2]Yale University, New Haven, CT

Methodological Question Being Addressed: Does alcohol self-administration change with positive and negative valence systems in accordance with a person's drinking history?

Introduction (Aims): The aim of the Research Domain Criteria (RDoC) project is to "develop, for research purposes, new ways of classifying mental disorders based on behavioral dimensions and neurobiological measures". It is essential that we have objective ways to measure and define these disorders across their stages of development. Alcohol Use Disorders are theorized to have 3 stages: binge-intoxication, preoccupation-anticipation, and withdrawal-negative affect. It is thought that the transition between these stages is sustained first through positive reinforcement, then negative reinforcement. The RDoC project refers to these as positive and negative valence systems. Particular to substance use disorders is self-administration behavior, a critical addiction phenotype to measure. Computer-Assisted Self-infusion of Ethanol (CASE) is a method that employs operant intravenous alcohol self-administration (IV-ASA) using a physiologically-based pharmacokinetic model-based algorithm that results in precise and well-controlled profiles of alcohol exposure across participants. Our aims were to (1) determine if recent drinking history is reflected in IV-ASA behavior using a free-access (FA) paradigm (2) to measure IV-ASA within a positive valence framework using a progressive-ratio (PR) paradigm (3) to assess how a negative valence system changes IV-ASA using guided imagery scripts to induce acute stress.

Methods: Healthy non-dependent drinkers (N=72) underwent an FA and PR self-administration session. Each session consisted of a priming phase of 4 individually standardized alcohol infusions, followed by a phase where they could push a button adlib to receive additional infusions. A PR session required participants to push a button an increasing number of times per infusion. A similar group (N=16) was stratified into binge/non-binge drinkers and participated in 3 FA sessions. Each session began with an exposure to a 5-min audio recording of personalized guided imagery scripts to induce acute stress, alcohol craving, and a neutral-relaxing state. This was followed by a 120-min FA session. Drinking history, impulsivity, and craving were measured using the Timeline Follow-back questionnaire, Delayed Discounting task, and the Alcohol Urge Questionnaire.

Results: Results demonstrated a strong association between recent drinking history, craving and IV-ASA during the FA paradigm. Those with heavier drinking histories, craving, and greater impulsivity scores had greater IV-ASA during the

PR session. Participants also showed greater IV-ASA following the stress-cue script in comparison to the neutral-cue script. Binge drinkers showed an approximately 2-fold higher IV-ASA following the stress-cue relative to neutral-cue, while non-binge drinkers did not show any difference across cue-type (all p's<0.05).

Conclusions: CASE is an effective and novel method to objectively measure IV-ASA in the laboratory. A positive valence system supports greater IV-ASA in those with heavier drinking histories and greater impulsivity, as shown in our PR session. We also were able to demonstrate that stress modulates IV-ASA and this relationship is mediated by drinking history. The CASE method and its paradigms provide objective measures that may help characterize the stage, severity, and risk for progression of alcohol addiction within the RDoC framework.

Disclosures: Supported by the NIAAA Division of Intramural Clinical and Biological Research (ZIA AA000466).

20 Applying Cue-Reactivity and Alcohol Self-Administration Procedures in a Bar-Like Laboratory to Explore Baclofen's Impact on Alcohol Drinking

Farokhnia M[1], Amodio JP[1], Bollinger JW[1], Edwards S[2], Zywiak WH[2,3], McGeary JE[2,4], Swift RM[2,4], Kenna GA[2], Farinelli LA[1], Leggio L[1,2]

[1]Section on Clinical Psychoneuroendocrinology and Neuropsychopharmacology, National Institute on Alcohol Abuse and Alcoholism (NIAAA) National Institute on Drug Abuse (NIDA) National Institutes of Health (NIH), Bethesda, MD, [2]Center for Alcohol and Addiction Studies, Brown University, Providence, RI, [3]Decision Sciences Institute, P. I. R. E., Pawtucket, RI, [4]Veterans Affairs Medical Center, Providence, RI

Methodological Question Being Addressed: Considering limited pharmacological treatments for alcoholism, there is a promising potential to test drugs that are already approved for other indications, and are also hypothesized to target alcohol-related neurobiological pathways. In this study, we combined two procedures (alcohol cue-reactivity and alcohol self-administration) to test the effects of baclofen on drinking behavior in alcohol dependent individuals. These are well-validated procedures, but rarely combined together. This design provides a comprehensive evaluation of the biobehavioral mechanisms by which medications affect alcohol drinking. Laboratory models have been found to be very informative and helpful in understanding the mechanisms of pharmacotherapy response, finding the proper population of responders, and finally optimizing the treatment of alcohol use disorder. In the present study, the laboratory experiment was conducted individually in a private bar-like testing room. The goal of using a bar-like room is to provide clinically relevant information that mirrors a real word scenario when people usually consume alcohol. This setting and the procedures provide a unique opportunity to test the effect of novel medications and collect accurate behavioral and biological markers of medication response.

Introduction: There is a substantial need to explore novel medications for alcohol use disorders. Previous preclinical and clinical work suggests that the GABAB receptor agonist baclofen may represent a promising pharmacotherapy for alcohol dependence (AD). However, more research is needed to better understand the biobehavioral mechanisms of baclofen and also to identify which alcoholic patients may respond to baclofen treatment.

Methods: This was a between-subject, double-blind, placebo-controlled, randomized human laboratory study. Fourteen non-treatment seeking AD individuals were randomized to receive baclofen (30 mg/day) or placebo. After being dosed for one week, participants underwent an alcohol laboratory experiment in an ecologically valid bar-like testing room that mirrors the real- world setting. This experiment was a combination of two established procedures: alcohol cue- reactivity (CR) followed by alcohol self-administration (ASA). CR: Participants were exposed to visual, tactile, olfactory, and proprioceptive stimuli associated with the beverage in a water trial followed by two consecutive alcohol trials. Cue-induced physiological changes (blood pressure, heart rate, salivation) as well as subjective measures of craving were recorded during each trial. ASA: A priming drink (designed to raise BrAC to approximately 0.03 g/dl) was first presented ("alcohol challenge" phase). This priming alcohol dose was used to assess the influence of resumption of alcohol drinking following a period of abstinence. Forty minutes later, more alcohol was presented as 8 mini-drinks (each designed to raise BrAC by 0.015%). Participants were allowed to drink from zero up to all mini-drinks ("free-choice" phase); $3 per drink was provided as an alternative reward for not drinking. Total amount of alcohol consumed as well as subjective effects of alcohol were evaluated during this session.

Results: During the CR session, there were no significant differences between the two groups in terms of urge and attention to alcohol (p>0.05). Baclofen-treated patients experienced significantly more salivation (p<0.01) and higher MAP (0=0.03) compared to the placebo group. During the ASA, participants who received baclofen consumed lower amounts of alcohol than the placebo group (0.17±0.41 vs. 1.43±2.30 standard drinking units, p=0.20). Although this difference was not statistically significant, there was a robust medication effect (d=0.76). When drinking during the ASA plus two days before was analyzed, there was a significant effect of baclofen to reduce alcohol consumption (p<0.01). After consuming

the priming drink, there was a significant effect of baclofen, compared to placebo, on the biphasic effects of alcohol, i.e. significant increases in alcohol stimulation ($p=0.001$) and sedation ($p<0.01$).

Conclusions: Baclofen reduced alcohol consumption during both the naturalistic phase and the laboratory experiment. The present study suggests that baclofen's effects on biphasic effects of alcohol (stimulation and sedation) might represent a potential mechanism of action in reducing alcohol drinking. Previous trials suggest AD patients with higher levels of anxiety may be particularly responsive to baclofen treatment. Our team is currently performing a follow-up larger study using a similar design but targeting specifically AD individuals with high anxiety levels.

21 A Phase 1 Single- and Multiple-Rising Dose Study of the Safety & PK of EMB-001, a Potential Treatment for Substance Use Disorders

Detke, MJ[1,2], Gloff, C[1], Connor, G[1], Straub, J[1], Robbins, A[1], Feltner, D[1] and Goeders, N[1,3]

[1]Embera NeuroTherapeutics, Inc; [2]Indiana University School of Medicine; [3]LSU Health Sciences Center

Methodological Question Being Addressed: The key methodological question was whether an innovative combination of SRD and MRD was appropriate to assess safety and PK for a combination product made up of two drugs, one of which had been approved for single-day use and the other for repeated dosing.

Introduction: EMB-001 is a combination of two FDA-approved drugs: metyrapone (MET), a cortisol synthesis inhibitor, and oxazepam (OX), a benzodiazepine. MET is approved for only one day of use as a test of pituitary function; OX is approved for acute and chronic treatment of various anxiety disorders. EMB-001 reduced cocaine and nicotine self-administration and attenuated cocaine and methamphetamine cue reactivity in rats. In a human study in cocaine-dependent subjects, EMB-001 significantly reduced cocaine use.

Methods: This was a single- and multiple-rising dose study; this innovative design was appropriate given the previous regulatory approvals and use of the individual drugs. Healthy volunteers, aged 18-65, received a single am dose on Day 1, BID dosing on Days 3-9 and a single am dose on Day 10. Three sequential dose cohorts of 8 subjects (6 drug, 2 placebo) received the following doses of MET and OX, respectively: 270 and 12 mg; 540 and 24 mg; and 720 and 24 mg. Total daily doses were double these amounts on BID dosing days. Primary outcomes were safety and the pharmacokinetics of MET, its active metabolite metyrapol, and OX. Safety measures included vital signs, ECGs and standard safety labs. Cortisol and other HPA axis parameters were monitored closely throughout the study.

Results: The most frequent adverse event was somnolence. Most AEs were mild; all were mild or moderate. There were no SAEs and no discontinuations due to AEs. Serum cortisol was reduced 2-4 hours after the first dose, consistent with the known pharmacology of MET, but had returned to baseline on subsequent mornings and at follow-up. There were no clinically significant changes in vital signs, ECGs or other safety labs. One subject in Dose Cohort 2 experienced a decrease in morning cortisol >50% relative to screening but was asymptomatic. Study drug was withheld for one day (Day 8), during which ACTH stimulation testing revealed sufficient adrenal response. Dosing was resumed and the subject completed the study. The half-lives of MET, OX and metyrapol were approximately and respectively 2, 7.5 and 8 hr. Exposure increased with increasing dose. There was modest accumulation with repeated dosing.

Conclusions: EMB-001 was well-tolerated in this study and no new safety signals were identified. These findings are generally consistent with MET and OX approved labeling and with safety data in 6 published studies, in which MET doses of 500-4000 mg/day were given for 2-8 weeks. PK results suggest that twice-daily dosing may provide appropriate duration of exposure for efficacy. Future plans include Phase 2 efficacy studies in cocaine use disorder and/or tobacco use disorder.

Disclosures: One or more authors report potential conflicts which are described in the program and as follows: All have been employees of Embera NeuroTherapeutics, Inc. which provided funding for this study.

22 Reduced Cortical Thickness in Combat-Exposed Veterans With and Without Experience of Early Life Abuse and Neglect

Averill LA,[1,2], Abdallah CG,[1,2], Pietrzak RH,[1,2], Southwick SM,[1,2], Levy I,[1,3], Harpaz-Rotem I[1,2]

[1]Clinical Neurosciences Division – National Center for PTSD, VA Connecticut Healthcare System, West Haven, CT, USA, [2]Yale University School of Medicine, Department of Psychiatry, New Haven, CT, USA, [3]Yale University School of Medicine, Department of Comparative Medicine and Neuroscience

Methodological Question Being Addressed: This study points to the use of multi-method biomedical imaging and clinical assessment to advance our understanding of the neurobiology of prolonged stress and trauma, to investigate and

identify potential biomarkers and treatment targets unique to specific types of trauma exposure, and to ultimately inform novel CNS drug development for improved treatment interventions.

Introduction: Early life trauma influences neural development and heightens risk of developing psychological disorders and vulnerability to negative outcomes following future exposure. Two recent studies have investigated cortical alterations in Veterans exposed to various traumas including early childhood and combat, both of which are highly pathogenic (Corbo et al., 2014; Woodward et al., 2009). These studies raise questions as to the direction and causality between these neural alterations and the development of PTSD. Do the differences in cortical integrity represent the consequences of early life trauma, combat exposure, the compounding effects of multiple traumatic experiences, or other factors? This pilot study extends the work of Corbo and Woodward to investigate the impact of combat on reductions in cortical thickness, controlling for age and childhood trauma.

Methods: Twenty-one combat-exposed Veterans with PTSD and 20 non-PTSD combat-exposed controls (mean age 32.7) completed the Combat Exposure Scale, Childhood Trauma Questionnaire and structural magnetic resonance imaging in a Siemens 3T TIM trio system. Image processing and analysis was completed in FreeSurfer. GLM was used to examine the effect of combat exposure on cortical thickness, controlling for early life trauma exposure and age using cluster-wise correction ($p<0.05$).

Results: Even after controlling for age and childhood trauma, and accounting for PTSD diagnosis, combat exposure is correlated to reductions in cortical thickness in the superior temporal ($z= -2.12$, $p < 0.05$) and rostral middle frontal regions ($z=-3.70$, $p < 0.05$). A negative correlation interaction between PTSD and combat was in the superior temporal/insular region ($z= -1.47$, $p < 0.05$) with a stronger negative correlation found in the combat control group. The interaction suggests increased combat exposure relates to more severe reductions in cortical thickness.

Conclusions: The results indicate combat exposure affects cortical structure beyond possible alterations due to early life trauma exposure, highlighting the need for careful assessment of combat exposure in returning Veterans. These brain regions are associated with memory, response suppression, and emotion dysregulation. The nature of combat, requiring heightened awareness of self and surroundings, continual sensory input, and regular manipulation/retrieval of data, may influence the demonstrated structural alterations. Multi-method assessment may provide insights into neural biomarkers and treatment targets that can advance our understanding of the neurobiology of chronic trauma exposure and inform novel CNS drug development.

Disclosures: All authors declare no conflict of interest related to this abstract.

23 Statistical Models Assessing Disease Modification Drug Effects Using Doubly Randomized Delayed-Start and Matched Control Designs

Turkoz I,[1] Sobel M,[2] Fu D-J,[3] Alphs L[3]

[1]Janssen Research and Development, LLC, Titusville, NJ, USA, [2]Temple University, Philadelphia, PA, USA, [3]Janssen Scientific Affairs, LLC, Titusville, NJ, USA

The Methodological Question Being Addressed: To evaluate how the study design along with analytical methods described here will have an impact on critical global public health issues and how they will address critical limitations of the randomized delayed-start design that were previously unresolved.

Introduction (Aims): Traditional randomized delayed-start designs for assessing disease modification have significant analytic limitations. This is because following an initial randomization, a high percentage of subjects in the control arm are likely to drop out before entering into the delayed-start period due to the long treatment period necessary to show the treatment's initial effects on disease progression. The proposed study design addresses limitations associated with alternative delayed-start trial designs that might be used to support a disease modification indication. Our proposed innovations represent hybrids of randomized and epidemiologic designs. We describe a design with a run-in period followed by a second randomization just prior to initiation of the delayed start. The run-in period is used to establish tolerability of the treatment and to provide a basis for a matched control rerandomization. The second randomization generates groups that are comparable with respect to disease characteristics that minimize confounding and bias.

Methods: The analytic objective of disease modification designs is to show that postponement of initiation of the treatment of interest does not permit the same ultimate level of recovery as that experienced by subjects who have been on the treatment from earlier points in their disease course. A number of models have been developed to identify "disease modification." The "failure-to-catch-up" concept is evaluated using Bayesian hierarchical models. Bayesian methodology

characterizes the truth of the failure-to-catch-up concept using posterior probabilities. This is a natural fit for establishing disease modification compared to traditional noninferiority margins or parallel-line approaches.

Results: Models to assess disease modification can be divided into two groups. The first group utilizes a growth curve describing disease evolution with slope parameters that are the same for all subjects. This group includes both simple linear and simple nonlinear spline models. The second group is based on a growth curve approach based on a subject-specific disease course evolution. Models in this latter group have both subject-specific and aggregated slope parameters that we evaluate using a hierarchical Bayesian structure. This second group of models includes linear random intercept models, linear random intercept plus random slope models, linear subject-specific spline models, and models that average over any or all of these. We compare the models with regard to how well they fit the current data and how well they can predict the future. The former is characterized by the deviance information criterion; the latter is characterized by Bayesian measures of fit.

Conclusions: Disease modification is a key goal for the treatment of chronic diseases (eg, schizophrenia, Alzheimer's disease, rheumatoid arthritis). The approach described addresses important limitations of alternative designs to establishing disease modification.

Disclosure: Support: Janssen Scientific Affairs, LLC. One or more authors report potential conflicts, which are described in the program.

24 Identifying Biomarker Signatures for Neurodegenerative Diseases using Large-Scale Network-structured Neuroimaging Measures

Li X[1], Zeng D[2], Marder K[1], Wang Y1

[1]Columbia University, [2]University of North Carolina at Chapel Hill

Methodological Question Being Addressed: Research on neuropsychiatric disorders faces unique challenges, including the fact that these disorders typically have a late onset and slow progression, and there is substantial disease and subject heterogeneity. In the proposed work, we aim to tackle these challenges by using subtle clinical measures, genomic markers, and neuroimaging biomarkers to develop statistical approaches to model the occurrence and progression of neurodegenerative disease and thereby inform the design of clinical trials.
Specifically, through regularization, we build a predictive model by taking into account the high dimensionality and network structure among the neuroimaging biomarkers, to yield an accurate disease risk prediction before manifesting clinical symptoms and biomarker signatures with better biological interpretation. With the predictive biomarkers identified by our model, they can be used to calculate sample size of pre-symptomatic subjects enrolled in the clinical trials for neurological disorders, or used as enrichment strategy through recruiting subjects at high risk and closer to clinical onset of a disease. The developed biomarker signatures can also be used to monitor disease progression. Lastly, estimating biomarker's association clinical diagnosis assesses their utility as surrogate markers in early phase clinical trials.

Introduction: Advances in biomarker data collection technologies provide exciting opportunities to study disease etiology and pathology, but at the same time pose challenges for statistical analyses. High throughput technologies generate high-dimensional large-scale biomarkers, some of which are involved in the same pathway of the underlying etiological mechanism and thus connected through biological networks. Moreover, the effects of biomarkers on the disease clinical onset in premanifest individuals may vary with a disease burden score (e.g., genetic risk factors) representative of a subject's disease susceptibility at the baseline. Biomarkers may exhibit large effects for subjects with high disease burden but not for subjects with low burden. However, which biomarkers manifest large effect for which groups of subjects is unknown.

Methods: To address these challenges in developing biomarker signatures for disease onset, we propose a varying-coefficient proportional hazards model to incorporate high-dimensional and network-structured biomarkers to predict time to disease onset. The regression coefficients are modelled fl unrestricted functions of the disease burden score. For estimation, we propose a doubly regularized local partial likelihood function, where a weighted L1-penalty is used to re-trieve sparsity in the presence of the high-dimensional covariates and a Laplacian regularization is used to incorporate the network information among the covariates. By assuming that the regression coefficients for two linked biomarkers share similar effect sizes locally at a given disease burden score, we reduce the chance of overfi and improve the reproducibility of discovered biomarker signatures in independent studies. The central distinguishing feature of the model is the nonparametric local network effect that captures the change of effect according to the disease burden score. We describe an efficient algorithm to estimate the regression coefficients, establish theoretical properties of the effect estimators, and derive weak oracle properties. Key advantages of the proposed method are the induced double smoothness over both the

dimension of the disease burden score and the space of network structure, the ability to incorporate non- linear interactions between high-dimensional biomarkers and the burden score, and the ability to achieve desirable model sparseness.

Results: Our numeric simulations demonstrate significant improvements over existing methods: in several settings, the proposed method has at least 3.5 times smaller integrated mean square error (IMSE: 2.18 - 3.17), compared with the commonly used non-local method under Lasso penalty (IMSE: 11.08 - 16.64). Regarding the variable selection properties, the proposed method has much improved false positive rate of 0.4% and a true positive rate of 95.1% as compared to the alternative method (false positive rate ranging from 9.4% - 28.6%, and true positive rate 82.6% - 98.3%). Our developed methods are applied to a newly completed comprehensive epidemiological study on Huntington's disease (HD), where the whole brain structural brain imaging network is used to predict the age-at-onset of HD. We obtained distinct biomarker signatures active for predicting HD onset with the effects depending on a subject's baseline disease burden score (e.g., produce of CAG repeats length and baseline age), and identified 7, 9 and 8 neuroimaging measures in Putamen and Thalamus regions most predictive for each group. The corresponding C-index was 0.96, 0.90 and 0.95 for low, median and high risk group- s, respectively, comparing with that of the non-local method (0.83). The results imply our method produced a sparse model with better predictive power.

Conclusion: The proposed methods estimate network biomarker signatures for a neurological disorder through identifying and combining important markers from a large candidate pool. They predict asymptomatic subjects who are at high risk of developing clinical onset within 5 years, and thus can be used in enrichment strategies to improve power for clinical trials by recruiting high risk premanifest subjects presenting certain signatures.

Disclosure: This work is supported by NIH grants NS073671, NS082062, NS036630, P01CA142538, P01CA142538, National Center for Research Resources grant UL1 RR025747.

25 Improved EEG-Based Biomarkers for Sleep Quality Based on Probabilistic Sleep Models

Dorffner G[1,2], Aydemir Ö[1,3]

[1]Medical University of Vienna, Austria, [2]The Siesta Group, Vienna, Austria, [3]Karadeniz Technical University, Trabzon, Turkey

Methodological Question Being Addressed: How can biomarkers for sleep based on EEG be improved with respect to their correlation to subjective sleep quality?

Introduction: Many CNS-active drugs, not just hypnotics, have a major effect on sleep that needs to be objectively described. The gold standard biomarker for measuring sleep is Electroencephalography (EEG), as part of polysomnography (PSG), together with other electrophysiological measurements. However, it has been known for a long time that correlations of PSG variables to subjective sleep quality assessments (from rating scales or questionnaires asking subjects how they slept) are very poor. In this study we explore novel probabilistic sleep models based on EEG data to prove that such correlations can be increased by extracting more information about sleep microstructure.

Methods: We applied a previously developed probabilistic sleep model (PSM, Lewandowski et al. 2012) to data from 270 subjects, both healthy sleepers and sleep-disturbed patients, and calculated proper descriptive variables (like area under the curve, entropy, spectral content, etc.) from the resulting probabilistic curves representing the main sleep processes such as light, deep or REM-like sleep. We then calculated linear regression models and optimized subsets of such variables regressing on a standard subjective sleep assessment score (subjective sleep and awakening assessment SSA). For comparison, we performed the same procedure with variables from traditional PSG-based sleep variables (such as sleep efficiency, percentage in each sleep stage, sleep latency, etc.).

Results: Results show that Pearson correlation coefficients between sets of probabilistic sleep variables and subjective sleep quality are significantly higher than for any corresponding set of traditional sleep variables (by 35% on average, e.g. from $r = 0.40 \pm 0.042$ for PSG to $r = 0.52 \pm 0.035$ for the PSM). We also found marked differences in sleep microstructure between individual subjects such that sleep profiles can be considered a kind of "fingerprint", pointing toward the distinctive power of the new biomarkers.

Conclusions: This work confirms that a EEG-based description of human sleep that goes beyond traditional sleep stages can extract more information about sleep and thus constitutes a better biomarker for measuring how well a person slept subjectively. This can be an important tool for early clinical trials on hypnotics and beyond, owing to the importance of sleep in the efficacy of CSN-active drugs.

Disclosure: Georg Dorffner is a part-time employee and shareholder of The Siesta Group GmbH.

Reference: Lewandowski A, Rosipal R, Dorffner G. 2012: Extracting more information from EEG recordings for a better description of sleep. Comput Methods Programs Biomed. 2012 Dec;108(3):961-72.

26 Precision of estimating the sample size needed to power clinical trials: What is the effect of risk-based monitoring?

Yavorsky C[1], Meares K[1], McNamara C[1], DiClemente G[1], Burger F[1]

[1]Cronos CCS

Methodological Question Being Addressed: Does risk-based data monitoring have an effect on power and sample size estimates in clinical trials?

Introduction: Determining the optimal sample size is a critical step in the design of a proposed clinical trial. Sample size calculations for clinical trials typically follow one of the following three approaches:

1) Specify the desired width of a confidence interval and determine the sample size; or
2) Bayesian approach that adjusts utility function; or
3) Analyzing the power of a test of hypothesis.

Enrolling too many subjects in a trial can be expensive and prolong the length of the trial. Likewise, if the study is not adequately powered, the statistical and clinical outcomes become questionable and may expose the subject to a potentially harmful treatment without enhancing knowledge. The current study proposes a novel method to sample size computation and utilizes a Bayesian approach that requires consideration of scientific goals, incorporating pilot data, study design, estimate of subject withdrawals, and estimate of data removed following risk based monitoring (RBM). Unlike current methods of sample size computation, Bayesian methods take into account the uncertainty of parameter estimates.

Methods: Data were evaluated retrospectively from a large multicenter randomized controlled trial for patients with schizophrenia. Power was calculated using the PAS software suite. Approximately 140 Site Raters participated in the study. A risk based monitoring paradigm was deployed to review psychometric and clinical scoring inconsistencies in PANSS data for within scale, within visit, across visits, within rater, and across raters. In addition to an automated data driven algorithm review, a an expert clinical review was conducted. Power and Effect Size computations utilizing both a Frequentist and Bayesian approach were completed for 4 groups as follows: 1) using prior data without dropouts, 2) using prior data with dropouts, 3) using prior data with RBM (with and without dropouts), 4) incorporating an unblinding sample size re-estimation design following RBM of 20% of assessments.

Results: Of the 2,361 PANSS assessments reviewed, 28.46% (n = 672) were "flagged" following RBM and clinical review, and 62.73% (n = 1481) had minimal to no discrepancies "passed." 1) using prior data without dropouts (as in common in schizophrenia trials), a total of 204 subjects were estimated, 2) using prior data with dropouts, a total of 340 subjects is needed as per the initial study, 3) a. using prior data with RBM dataset without 40% drop out rates, 170 subjects are needed, b. using RBM dataset with 40% drop out rates, 238 subjects are needed, 4) incorporating an unblinding sample size re-estimation design, 159 subjects are needed without the 40% dropout rate and 223 subjects.

Conclusions: In all cases utilizing RBM, the sample size needed for hypothesis testing was reduced from the original by at least 30%, with the unblinding sample size re-estimation following RBM showing the most significant reduction in the number of subjects needed. Inadequately sized studies often results in investigator's unrealistic assumptions about the effectiveness of study treatment. Using an adequate sample size with high quality data collection through RBM could result in more reliable results, more timely trials and fewer resources.

27 Human Factors Evaluation of a Novel Digital Health Feedback System in Psychiatry

Peters-Strickland T[1], Smith JL[2], Bartfeld B[1], Pestreich L[1], Rohatagi S[1], Hatch A[1], Forma F[3], Raja P[4], DochertyJ[1]

[1]Otsuka Pharmaceutical Development & Commercialization, Inc., Princeton, NJ, [2]GfK Custom Research LLC, Oakbrook Terrace, IL, [3]Otsuka America Pharmaceutical, Inc., Princeton, NJ, [4]Proteus Digital Health, Inc., Redwood City, CA

Methodological Question Being Addressed: Poor medication adherence is a common problem in patients with serious mental illness (SMI). The Digital Health Feedback System (DHFS) is a new class of combination drug-device developed for patients with SMI comprising a sensor-embedded medication, wearable sensor, and software applications. The DHFS offers a new opportunity to objectively measure patient's medication ingestion and share this data with healthcare professionals (HCPs). The absence of directly applicable user experience from a comparable existing product highlights

the importance of applying human factors (HF) methods to analyze use-related risks and optimize the system. HF Engineering is recommended in the FDA guidance to assess the safe and effective use of a system by the intended users for the intended uses. This is particularly critical for patients with SMI who may have cognitive impairments associated with poor functional skills.

Introduction (Aims): To design the DHFS to be safe and effective by using HF methodology to identify steps in the use process that may result in use-related risks, understand the root cause of performance problems, and iteratively mitigate risks by redesigning the product.
Methods: Three successive formative HF studies were conducted. Each study tested users in the intended user group which included patients with schizophrenia, major depressive disorder, and bipolar1 disorder. Both objective and subjective qualitative data was gathered for all tasks considered critical to safety or essential for effective use. The root cause for errors was assessed and risks to the users were identified. A use risk analysis was performed before and after each study to identify remaining user risks associated with the system and inform iterations of product design. With this methodology, use-related hazards were iteratively mitigated throughout the design process, while also improving effective use by the
intended users.

Results: Feedback from the formative HF studies of the patient interface was used to implement further design modifications to the patient app and electronic instructions for use. The modifications are in line with design characteristics of digital health applications that reduce the cognitive effort needed to effectively use an app for persons with SMI, including: minimizing the number of the application levels (hierarchy), simplifying content (sentences, composition, reading level), using explicit wording, and avoiding information overload. There were no distinguishable differences in performance observed among patients with different diagnoses. Risk analysis performed after a final validation study indicated that the mitigation effort eliminated potential failure modes that might lead to unacceptable risks.

Conclusions: The results from iterative studies of the patient interface of the DHFS results in a system that is designed to be safe and effective for its intended users, intended uses, and use environments.

Disclosure: Otsuka Pharmaceutical Development & Commercialization, Inc.

28 Objective Measurement of Compliance and Analysis of Per Protocol Data Set in Clinical Trials of Tardive Dyskinesia

O'Brien CF, Burke J, Jimenez R, Siegert S, Aurora B

Neurocrine Biosciences, Inc., San Diego, CA

Methodological Question Being Addressed: To what degree does noncompliance exist in neuropsychiatric studies and how might sponsors account for it?

Introduction: Common methods to measure compliance with interventions in clinical trials include patient feedback and "pill" counts. Although pill counts are an attractive option given the simplicity of the approach, accuracy of this method is dependent upon the patient, who may be tempted to discard pills to appear more compliant. One may assume that clinical trials would have a higher compliance rate compared with routine clinical practice given the level of patient attention/follow-up; however, clinical trials report compliance rates of only 43%-78% among patients treating chronic conditions. A lack of compliance represents a considerable source of methodological bias and could ultimately jeopardize clinical trial results.

Methods: Two randomized, double-blind, placebo-controlled trials have investigated the safety and efficacy of once-daily valbenazine (VBZ), a novel, highly selective vesicular monoamine transporter 2 inhibitor with a 20-hour half-life, for treatment of tardive dyskinesia in subjects with schizophrenia, schizoaffective disorder and mood disorder. KINECT 2 (NCT01733121) evaluated VBZ 25-75mg/day, while KINECT 3 (NCT02274558) evaluated VBZ 40 or 80mg/day. Primary endpoints for both studies were analyses of the Abnormal Involuntary Movement Scale (AIMS) score for the intent-to-treat (ITT) population at Week 6, which included all subjects with baseline and ≥1 post-baseline AIMS. The per protocol population was defined *a priori* for supportive analyses and required, for subjects randomized to VBZ, detectable plasma VBZ at Week 6. Subjects in both trials were to return unused study drug and empty drug packaging material to the study center at each visit. Drug compliance was measured by qualified personnel counting capsules returned.

Results: 176 subjects receiving VBZ (n=44 KINECT 2, n=132 KINECT 3) were included in the ITT analysis and had a Week 6 pharmacokinetic assessment. Of these, 11 (25.0%) VBZ subjects in KINECT 2 and 15 (11.4%) VBZ subjects in KINECT 3 had below quantifiable levels (BQL) of VBZ at Week 6. Seven VBZ subjects in KINECT 2 and 6 VBZ subjects

in KINECT 3 had BQL at all three study visits, Weeks 2, 4, and 6. When evaluating compliance by pill count, 10/11 KINECT 2 and 13/15 KINECT 3 subjects reported taking all doses throughout the six-week double-blind treatment phase and returned empty study drug packaging material. These observations were generally consistent regardless of age, gender, diagnosis group, VBZ dose, or psychiatric status at screening. Non-compliance was not due to adverse events, and BQL concentrations were accurately assayed.

Conclusions: Consistent with what has been reported previously, there was discordance between actual exposure compliance rates and pill count compliance rates. All subjects deemed noncompliant based on plasma levels returned empty or nearly empty packaging material, suggesting that they tampered with drug packaging. Objective, pharmacokinetic assessment of compliance is more accurate than counts of returned medication and, in conjunction with per protocol analyses, appears to be justified in clinical trials of subjects with underlying psychiatric conditions. Further analysis of the impact of patient non-compliance on clinical trial results is warranted.

29 Obtaining a More Comprehensive Understanding of Site and Subject Preferences for Clinical Assessment Technologies: Key Implications for the Industry

Lytle D, Cohen EA, Rock CR, Komorowsky AE, Friedmann B, Murphy MF

Worldwide Clinical Trials, King of Prussia, PA

Methodological Question Being Addressed: How do sites and subjects view the use of clinical assessment technologies compared to traditional paper measures; how could the industry move forward in order to effectively implement these assessment solutions?

Introduction: As technology continues to grow, so does its use within the clinical trial industry. In particular, tablet- and computer-based eSource/e-Clinical Outcome Assessments (eCOA) and electronic Patient Reported Outcomes (ePRO) solutions are increasingly being used. In previously reported survey data related to international sites' preferences using eCOA solutions (Lytle et al., 2015), results indicated that sites have strong preferences for paper based assessments (paper assessments had a +62 basis-point preference ratio). Additional data were collected in the Lytle et al. global survey and is presented here to provide a more comprehensive perspective on site (and perceived subject's) preferences for electronic outcome solutions.

Methods: Site coordinators and raters were anonymously surveyed from US and ROW sites. The sites designated to receive the survey had previously participated in numerous psychiatric and neurocognition studies. Sites were queried in regards to the burden of eCOA solutions and paper-based assessments; we also obtained sites' preferences regarding the modalities of how the raters receive feedback from rater reliability groups. Finally, responses were obtained from sites regarding patient and caregiver feedback they received about the ePRO and paper based assessments.

Results: Data indicate that sites' preferences for paper based outcome measures over eCOA can partially be explained by the burden of preparing the latter for each patient (60% of respondents indicated that the preparation of the eCOA solution was cumbersome). Surveyed sites also reported that the eCOA device also presented a "barrier" or distraction during the patient interview. When sites do not use eCOA solutions for a particular study, they reported that the transmission of the traditional paper-based assessments was actually more of a burden than the preparation of eCOA solutions (68% of respondents indicated that this was cumbersome, with 75% of coordinators indicating that this was cumbersome); nevertheless, the sites continued to prefer paper assessments overall as compared to eCOA solutions. Sites also revealed interesting preferences for how they receive clinical feedback from a rater reliability group, with e-mail being the most favored mechanism (+81 basis-point preference ratio). Finally, the sites reported that patients and caregivers also strongly preferred paper based PROs (+70 basis-point preference ratio).

Conclusions: Understanding site preferences - and adapting to sites' issues in relation to paper based and electronic outcome assessments - will further the implementation of electronic solutions and ultimately enhance the overall quality of rater assessments. The poster provides suggestions of how the industry should be mindful of the findings in this study when developing and providing sites with eCOA solutions.

Disclosures: All authors are employees of Worldwide Clinical Trials and have no conflicts of interest.

30 Using κ' to Provide a Common Frame of Reference for Rater Agreement

Lord-Bessen J, Rothman B, Tatsumi K, Opler M

Prophase, LLC

Methodological Question Being Addressed: How can we better measure rater agreement when we use measures with nominal and ordinal categories?

Introduction (Aims): When Clinician Reported Outcome (ClinRO) measures that use nominal or ordinal rating categories are employed in CNS clinical trials, Cohen's kappa (κ) is often used as the measure of rater agreement. However, researchers are often unaware that the marginal distributions of the contingency table generated by two raters affect the upper and lower bounds of κ. By rescaling the components of κ—expected and observed agreement—a measure Cohen referred to as κ' can be created which allows for a common scale for tables with different marginal distributions. While κ has been a preferred measure for agreement, we demonstrate through examples of contingency tables differences between κ and κ' and the advantages of using the latter.

Methods: Using Karelitz & Budescu (2013) as a template, we simulate various examples of rater agreement tables under conditions that allow for comparison of varying marginal distributions. We calculate a measure of similarity between the row and column marginal referred to as Distance. We then calculate κ, the possible upper and lower bound based on the marginal, and the rescaled version of κ, κ'. Finally, we explore the comparisons to highlight the interpretational differences.

Results: For a 30-item measure with 3 ordinal categories, we generate three contingency tables. The first has uniform row and column margins of (10, 10, 10). Since they are identical, Distance = 0. This table results in a κ=0.55 with the upper and lower bounds of [0,1]. Because κ the full range of κ, κ'= κ. However, a table where the marginals vary in opposite directions (row marginal = (7,10,13); column marginal = (13,10,7); Distance = 0.4), the resulting κ=.56 is now bounded by an upper limit of 0.8. Rescaling κ results in κ'=0.79. A third table where the marginals vary in the same direction but are not identical (row marginal = (7,12,11); column marginal = (9,9,12) ; Distance = 0.2), the resulting κ=.55 is now bounded by an upper limit of 0.9. Rescaling κ results in κ'=0.64.

Conclusions: In all three cases, the estimate of κ was approximately equal, but κ' differentiates the three by acknowledging that the three scores are on a different range and provides a common frame of reference for comparison. This investigation highlights the importance of considering both the marginal and not just the joint agreement as they are two distinct components.

Disclosures: All authors are employees of ProPhase, LLC.

References: Karelitz, T.M. & Budescu, D.V. (2013). The effect of the raters' marginal distributions on their matched agreement: A rescaling framework for interpreting kappa. *Multivariate Behavioral Research, 48*(6), 923-952.