MODELING REACTION TIME DISTRIBUTIONS INCREASES THE STATISTICAL POWER OF COGNITION TESTING

Seth C. Hopkins,[1] Sean D. Wilson,[2] Ajay Ogirala,[1] Gabriel Stellman,[3] Snezana M. Milanovic,[1] Steven T. Szabo,[1] James A. Rogers,[2] Chris J. Edgar,[3] Kenneth S. Koblan[1]
[1]Sunovion Pharmaceuticals Inc; [2]Metrum Research Group, Tariffville, CT; [3]Cogstate Ltd, New Haven, CT

ABSTRACT

BACKGROUND: Drug development clinical trials often include tests of cognition to assess participants' cognitive performance during individual testing sessions. The testing sessions, designed to measure cognitive domains (such as psychomotor function, attention, visual learning and working memory) collect subjects' response data as reaction time (RT) and accuracy, and report subject-level metrics (eg, mean log-RT) that are used to quantify cognition. We examined whether current approaches using mean log-RT to reduce an individual subject's performance data results in information loss relative to other approaches that more fully model individual subject's RT distributions, and thereby reduce the power to test specific hypotheses of cognition.
METHODS: Reaction times were extracted from subject-level performance data during computerized cognitive tests (Cogstate tests of Detection, Identification, One Card Learning, One Back). Data from 7 drug development clinical trials (schizophrenia, bipolar depression, N=1,890 subjects) were compared to normative data obtained from healthy subjects (N=7,108). Parameters describing subject-level RT distributions were obtained by Bayesian estimation of population models using either ex-Gaussian or Wiener diffusion model residual likelihoods. Information loss was examined by comparing the ability of single parameter mean log RT to reject null hypotheses of cognition versus the parameters of RT distribution models (ex-Gaussian, Wiener Diffusion Model). Here we evaluated the sensitivity and specificity of the discrimination (AUC) between subjects with schizophrenia or bipolar depression versus healthy subjects.
RESULTS: An individual subject participating in a cognitive test session performed approximately 170-180 responses across the 4 tests. Subject-level RT distributions were well-described by ex-Gaussian and Wiener diffusion models, resulting in parameter estimates for each subject. The sensitivity/specificity of cognitive performance data alone to classify adults with schizophrenia or bipolar depression from healthy subjects was improved for each task. For example, correctly categorizing disease status was improved for the diffusion model versus mean log-RT, with AUC values of 81% vs. 77% for Identification, 78% vs. 69% for One Card Learning, and 74% vs. 62% for Detection.
DISCUSSION: Analyzing subject-level responses during cognitive testing recovers information lost by mean log-RT, the latter being most typically used in analyses of cognitive performance data. The ability to separate individuals with schizophrenia or bipolar depression from healthy controls using cognitive domains was improved by 4-12 percentage points across cognition tasks. In conclusion, modeling subject-level RT distributions is superior to the typical use of single performance metrics and improved analysis methods may increase the statistical power to test specific hypotheses of cognition in clinical trials.

BACKGROUND

- Drug development clinical trials often include tests of cognition to assess participants' cognitive performance during individual testing sessions
- The testing sessions, designed to measure cognitive domains (such as psychomotor function, attention, visual learning and working memory) collect subjects' response data as reaction time (RT) and accuracy, and report subject-level metrics (eg, mean log-RT) that are used to quantify cognitive performance
- Over the past 15-20 years, Bayesian estimates using either ex-Gaussian or Wiener diffusion model residual likelihoods have become increasingly influential as a methodology for modeling the psychological and neural aspects of processes involved in cognitive decision making
- We examined whether a commonly used analytic approach, using mean log-RT, results in information loss relative to other approaches that more fully model individual subject's RT distributions

OBJECTIVE

- To test whether RT distribution models (ex-Gaussian and Wiener Diffusion Models), as compared to single parameter (mean log-RT), are better able to discriminate performance on the Cogstate in subjects with a diagnosis of schizophrenia or bipolar depression versus healthy controls

METHODS

- Reaction times (RT) were extracted from subject-level performance data during the Cogstate battery, a widely used computerized cognitive test
- Cogstate tests consisted of Detection (assessing psychomotor function), Identification (assessing attention), One Card Learning (assessing visual learning), One Back (assessing working memory); Note: the One Back task was not included in the current investigation because no healthy control data were available for this task)
- Because the Cogstate Detection task has a single component (reaction time), the diffusion model is not identifiable, therefore the shifted Wald diffusion model was utilized for this task
- Baseline Cogstate data from 7 clinical trials in patients with a diagnosis of schizophrenia or bipolar depression were compared to normative data obtained from healthy subjects

Supported by Funding from Sunovion Pharmaceuticals Inc.
American College of Neuropsychopharmacology; December 5–8, 2021
Puerto Rico Convention Center, San Juan, Puerto Rico

METHODS (cont.)

- Parameters describing subject-level RT distributions were obtained by Bayesian estimation of population models using Wiener diffusion model residual likelihoods
- We evaluated the sensitivity and specificity of the 2 analytic methods for categorizing subjects with a diagnosis of schizophrenia or bipolar depression versus healthy subjects:
  (1) single parameter (mean log RT) versus
  (2) the parameters of RT distribution models (ex-Gaussian, Wiener Diffusion models; and Wald diffusion model for the Detection task)
- Specifically, receiver operating characteristic (ROC) curves were used to evaluate the performance characteristics (sensitivity, specificity) of each analytic method for detecting a difference in Cogstate function between the clinical group and the healthy control group
- Area Under the ROC Curve (AUC_{ROC}) was reported to summarize overall accuracy of each analytic approach

Wiener Diffusion Model: Example of Bayesian-derived parameters characterizing subject-level RT distributions

- The response time distribution is characterized by 4 main parameters (see **Fig. 1**, below):
- z = initial response bias; the subject's a priori expectation/bias about the upcoming stimulus
- α= response caution threshold (how "conservative" the subject is in deciding on their response)
- ν = drift parameter (a subject's information processing/cognitive speed)
- t_O = the non-decision-related parameters: decoding the stimulus, motor response time

Figure 1. Reaction Time Distributions Derived During Cognition Testing in Drug Development Clinical Trials

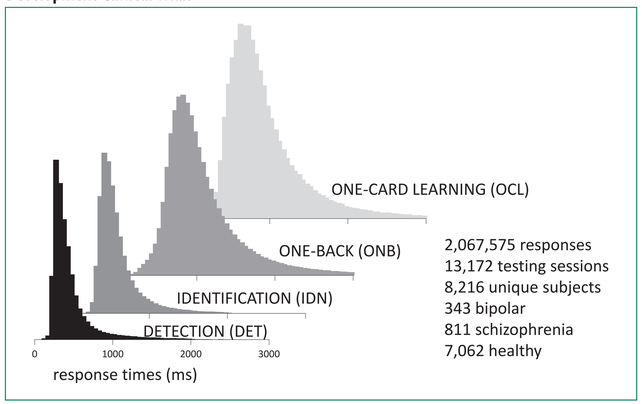
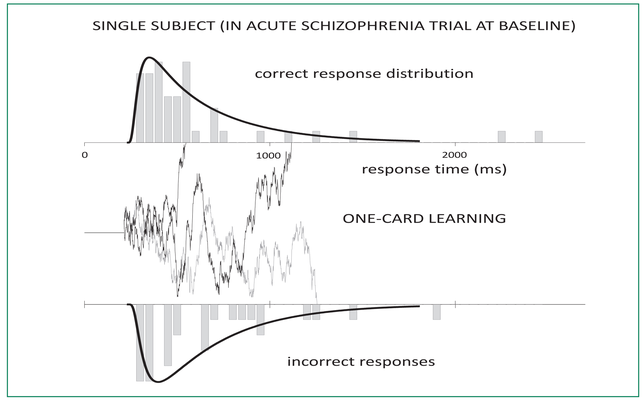


Figure 2. Fitting a Diffusion Model to Cognition Test Results in Drug Development Clinical Trials



RESULTS

Figure 3. Bayesian Hierarchical Fitting for Reaction Time Distributions: Example of a Single Subject Performing 4 Cognitive Tasks

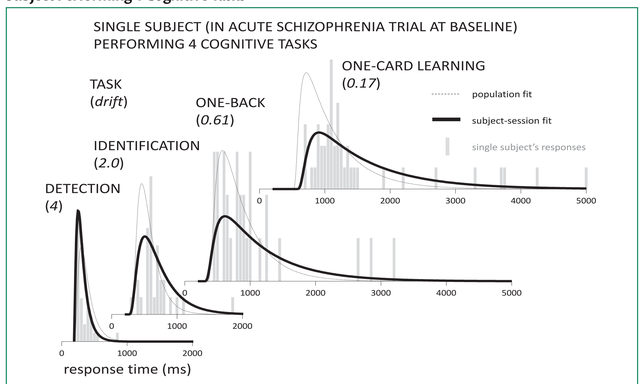


Figure 4: ROC Curves for Cogstate Detection, Identification, and One-Card Learning Tasks

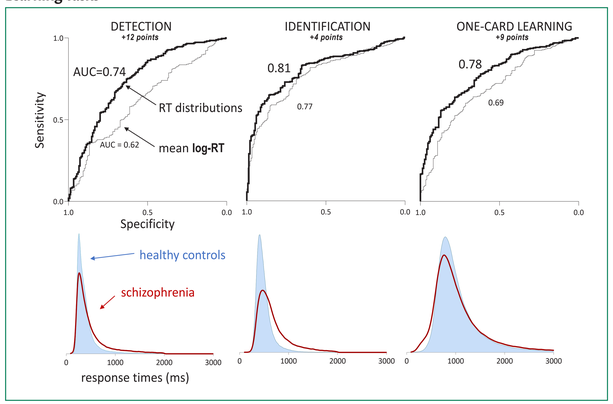
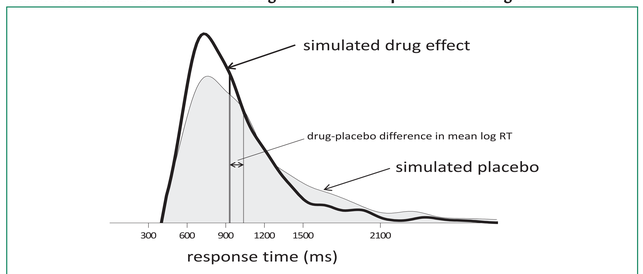


Table 1. Estimates of Drift Parameters Obtained From Fitting Diffusion Models in Clinical Trial Cognitive Testing

				COGSTATE TEST								PANSS TOTAL	
				DETECTION		IDENTIFICATION		ONE-CARD LEARNING		ONE -BACK			
ABSOLUTE VALUES AT BASELINE			STUDY	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
	Healthy	adults	CogState	5.1 (2.3)	350	3.5 (0.76)	350	0.56 (0.31)	349				
	Bipolar	adolescent	D1050326	3.3 (1.3)	340	2.8 (0.94)	341	0.57 (0.41)	341	1.5 (0.65)	341		
	Schizophrenia	adolescent	D1050301	3.4 (1.5)	317	2.7 (0.98)	324	0.48 (0.39)	324	1.4 (0.67)	324	94 (11)	326
		adult	D1050233			2.6 (1.1)	222	0.37 (0.34)	222	1.2 (0.76)	222	98 (11)	488
		adult <40y	SEP361201	4.0 (1.8)	210	3.1 (1.1)	214	0.56 (0.40)	214	1.6 (0.69)	214	101 (8.1)	245
EFFECT OF ULOTARONT/PLACEBO TREATMENT													
Baseline-to-Week-4	**Schizophrenia**	adult <40yr	SEP361201	-0.06 (1.8)	208	-0.02 (0.98)	214	0.00 (0.32)	214	0.04 (0.06)	214	-17 (15)	196
Baseline-to-Month-6			SEP361201	-0.1 (2.0)	119	0.05 (1.1)	123	0.09 (0.38)	123	0.09 (0.63)	123	-42 (14)	104

Figure 5. In Simulated Trials of a Drug Having an Underlying Drift Effect Size of 1 (Drug vs. Placebo): The Drift Parameter Demonstrated a Larger Effect Size Compared to Mean Log-RT



100% of trials detect statistically significant difference in drift parameter
14% detect difference in mean log-RT

Average effect size among drug vs. placebo trials:
Drift parameter = **0.86**
Mean log-RT parameter = **0.17**

DISCUSSION

- Analyzing subject-level responses during cognitive testing recovers information lost by mean log-RT, the latter being a widely used method for analyzing cognitive performance data
- The ability to correctly classify subjects with schizophrenia or bipolar depression (versus healthy controls) was improved by 4-12 percentage points using ex-Gaussian or Wiener diffusion models of baseline Cogstate data
- In conclusion, modeling subject-level RT distributions is superior to the typical use of single performance metrics and improved analysis methods may increase the statistical power to test specific hypotheses of cognition in clinical trials

DISCLOSURES

Drs. Hopkins, Ogirala, Milanovic, Szabo, and Koblan are employees of Sunovion Pharmaceuticals Inc.
Drs. Wilson and Rogers are employees of Metrum Research Group, working under a contract with Sunovion Pharmaceuticals Inc.
Drs. Stellman and Edgar are employees of Cogstate Ltd, working under a contract with Sunovion Pharmaceuticals Inc.
Medical writing support was provided by Edward Schweizer, MD of Paladin Consulting Group, and was funded by Sunovion Pharmaceuticals Inc.
Sunovion discovered ulotaront in collaboration with PsychoGenics based in part on a mechanism-independent approach using the in vivo phenotypic SmartCube® platform and associated artificial intelligence algorithms.