Relating Memory to Functional Performance in Normal Aging to Dementia Using Hierarchical Bayesian Cognitive Processing Models

William R. Shankle, MS, MD, FACP,†‡§ James P. Pooley, BS,§ Mark Steyvers, PhD,‡ Junko Hara, PhD,*‡ Tushar Mangrola, MS,‡ Barry Reisberg, MD,‖ and Michael D. Lee, PhD‡*

Abstract: Determining how cognition affects functional abilities is important in Alzheimer disease and related disorders. A total of 280 patients (normal or Alzheimer disease and related disorders) received a total of 1514 assessments using the functional assessment staging test (FAST) procedure and the MCI Screen. A hierarchical Bayesian cognitive processing model was created by embedding a signal detection theory model of the MCI Screen-delayed recognition memory task into a hierarchical Bayesian framework. The signal detection theory model used latent parameters of *discriminability* (memory process) and *response bias* (executive function) to predict, simultaneously, recognition memory performance for each patient and each FAST severity group. The *observed* recognition memory data did not distinguish the 6 FAST severity stages, but the latent parameters completely separated them. The latent parameters were also used successfully to transform the ordinal FAST measure into a continuous measure reflecting the underlying continuum of functional severity. Hierarchical Bayesian cognitive processing models applied to recognition memory data from clinical practice settings accurately translated a latent measure of cognition into a continuous measure of functional severity for both individuals and FAST groups. Such a translation links 2 levels of brain information processing and may enable more accurate correlations with other levels, such as those characterized by biomarkers.

Key Words: neuropsychological testing, wordlist memory, recognition memory, functional assessment staging test, clinical dementia rating

(*Alzheimer Dis Assoc Disord* 2012;00:000–000)

Received for publication June 7, 2011; accepted December 31, 2011.

From the *Shankle Clinic; †Memory and Cognitive Disorders Program, Hoag Neurosciences Institute, Newport Beach; ‡Medical Care Corporation; §Department of Cognitive Sciences, UC Irvine, Irvine, CA; and ‖Department of Psychiatry, New York University School of Medicine, New York, NY.

Supported by award NIRG-08-90460 from the Alzheimer's Association to M.D.L. and W.R.S., and by US Department of Health and Human Services grants AG03051 and AG08051, the Fisher Center for Alzheimer's Research Foundation, the Litwin Fund for Alzheimer's Disease Research, the Woodbourne Foundation, the Louis J. Kay and June E. Kay Foundation, the Stringer Foundation, and the Hagedorn Fund to B.R. No sponsors have roles in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Disclosure: W.R.S. is an employee of and a cotrustee of the trust holding the shares of Medical Care Corporation whose cognitive assessment tool (ie, MCI Screen) is used in this study, and servers as a paid speaker for Forest Pharmaceuticals. M.S. and M.D.L. serve on the board of scientific advisers of Medical Care Corporation. There is no conflict of interest issue that arises as a result of this position. J.H. is an employee of and a cotrustee of the trust holding the shares of Medical Care Corporation. T.M. is an employee of Medical Care Corporation and holds shares in Medical Care Corporation. B.R. is the developer and the copyright holder of the Functional Assessment Staging scale (FAST) that is used in this study. J.P.P. has nothing to disclose.

Reprints: Junko Hara, PhD, Shankle Clinic, 3900 West Coast Hwy, Ste 310, Newport Beach, CA 92663 (e-mail: junkoh@shankleclinic.com).

Relating cognitive to functional impairment has been a relatively understudied area in Alzheimer disease (AD) research. Yet, it is practically important in terms of understanding outcomes in clinical trials and in predicting the degree of impairment in functional capabilities from objective cognitive testing in clinical practice.

The usual way of relating cognition to function has been to look at their intercorrelations. An alternative approach is to create a model that posits the form of the relationship between the degree of impairment in functional capabilities and the processes underlying a given cognitive task. Recently, we applied a methodology combining hierarchical Bayesian statistical methods with psychological measurement models of the processes underlying memory [hierarchical Bayesian cognitive processing, (HBCP)]. Such models may provide useful insights into the cognitive ability being studied. They can also simultaneously estimate parameters for groups and individuals, automatically make inferences for missing data, and integrate multidimensional data, such as biomarkers, cognitive, and functional measures plus covariates, into a single construct.

The delayed recognition memory task may help relate cognitive and functional changes because its performance requires memory storage and executive function processes. The task involves studying a list of items (usually words or pictures) one or more times, and then after a few minutes or longer, presenting these *studied* (old) items intermixed with a list of *nonstudied* (new) items. The subject is asked to discriminate the *old* from the *new* items.

Signal detection theory (SDT) is used to model recognition memory as composed of underlying memory and decision-making processes.[1–3] Decision making is an executive function, which helps individuals perform various functional abilities.

The functional assessment staging procedure (FAST) is a valid and internationally used measure of the degree of impairment in functional capabilities for persons with AD, in which the functional stages have been correlated with cognitive impairment.[4,5] Because of their use in clinical practice and research, it is useful to explore how different degrees of functional incapacity relate to an SDT model of delayed recognition memory using clinical data.

METHODS

FAST Staging Procedure

At each patient visit, a trained physician interviewed either the patient or a reliable informant using the FAST procedure[5,6] to stage the patient's degree of functional incapacity into one of the 16 stages (7 major stages, 1 to 7, with 11 substages, 6a to e and 7a to f).[7]

Patients with no subjective functional impairment and no objectively evident functional impairment are classified as FAST stage 1 and will be referred to as no cognitively related functional impairment (NCI). Patients who subjectively have greater difficulty in cognitively related functional abilities but still perform them completely normally are classified as FAST stage 2. FAST stage 2 patients will be referred to as subjective cognitively related functional impairment (SCI). Patients who have impairment in cognitively related, executive-level functional abilities, such as using a calendar to prospectively keep appointments but have no impairment in instrumental activities of daily living (complex activities of daily life, such as marketing properly, managing personal finances properly, and preparing meals for guests, in one's accustomed manner), are classified as FAST stage 3. FAST stage 3 patients will be referred to as mild cognitively related functional impairment (MCI). FAST stage 4, 5, and 6 patients have functional deficits that correspond to the levels of mild, moderate, and moderately severe dementia, respectively, and are classified by their degree of impairment in instrumental and basic activities of daily living.

Cognitive Testing

At each visit, patients are tested with a cognitive battery derived from the Consortium to Establish a Registry for AD (CERAD), consisting of trails A and B—measures of sequencing, processing speed, and set shifting; FAS letter fluency—a measure of phonemic fluency, working memory, and rule application; Ishihara number naming test—a measure of object recognition that minimizes the use of semantic memory; CERAD drawings—a measure of simple object recognition, planning, organization, and visual constructional praxis; and the MCI Screen (MCIS)—a measure of rule application, working memory, rehearsed delayed recall and recognition memory, unrehearsed delayed recall, judged comparisons, and self-estimation of memory ability.[8–10] The MCIS was the only cognitive test involved in relating cognition to function.

Clinical Sample

The data came from a primary care and a cognitive disorder clinic and included all assessments of 280 patients followed up every 3 to 6 months for up to 6 years. The number of patients assessed one or more times in any given FAST stage varied from 26 to 163. If one sums the number of patients assessed per FAST stage over the 6 stages, the total is 514, which is greater than the 280 patients in the study. This is because each patient can contribute data to one or more FAST stages. There was a total of 1514 FAST stage assessments for the 280 patients studied.

Because the data are repeated measures, there are potential confounding effects on task performance due to practice and reliability. However, these potential confounds have been shown to be small and are therefore unlikely to influence the present study's results (MCIS interrater and test-retest reliability = 0.83[9]; wordlist effect size < 0.009 SDs[11]).

Patients with AD or a related disorder (ADRD) underwent a standardized evaluation, including MRI, laboratory tests, medical history, and physical examination and were diagnosed according to the published criteria for AD, Lewy body disease, cerebrovascular disease, and frontal temporal lobe disease.[12–16] Patients were followed up every 3 to 6 months from 2002 to 2007. Table 1 shows the numbers of patients and patient assessments by FAST stage, along with its description.

Delayed Recognition Memory Task

The delayed recognition memory task was performed after the MCIS-delayed free recall task. The examiner reads aloud to the patient the 10 study list words (*old*) intermixed with 10 unstudied words (*new*), one at a time, and the patient is asked to decide whether the word was *old* or *new*. In SDT, correct identifications of *old* and *new* words are called *hits* and *correct rejections*; incorrect identifications of *old* and *new* words are called *misses* and *false alarms*.

TABLE 1. Number of Study Patients and Patient Assessments in Each FAST Stage

	Patients		Assessments		FAST Stage*	
FAST Stage	N	%	N	%	Severity	Description
1	159	15%	288	19%	Normal	No subjective or objective, cognitively related functional decline
2	163	10%	308	20%	Normal/SCI	Subjective decline in cognitively related functional capacity
3	26	36%	129	9%	MCI	Objective impairment in cognitively related complex functions without impairment in instrumental ADLs
4	75	28%	436	29%	Mild dementia	Impaired instrumental ADLs
5	47	7%	189	12%	Moderate dementia	Impaired judgment related to proper selection of clothing for social and weather conditions
6	44	4%	164	11%	Moderately severe dementia	Impaired basic ADLs
All	514	100%	1514	100%		

ADL indicates activities of daily living; *FAST, functional assessment staging test, copyright©1984 by Barry Reisberg, MD; MCI, mild cognitively related functional impairment; SCI, subjective cognitively related functional impairment.

Construction of New Wordlists for the MCIS

Ten pairs of equivalent wordlists have been constructed to minimize practice effects, minimize interitem associability, and parallel the original CERAD wordlist. Eight of the 10 pairs are used with the MCIS test. Each time a patient is tested, the MCIS algorithm randomly selects a pair of *old* and *new* wordlists from the available pool without replacement. This means that the patient has to take the MCIS test 9 times before being exposed to the same pair of wordlists. The wordlists were designed so that: (1) the items of the *old* and *new* wordlists are similar; (2) words are 1 or 2 syllables; (3) their frequency, range, and diversity statistics resemble those of the original CERAD wordlist; (4) the words in each list are not easily associable (low semantic associability)[17]; (5) the residual semantic similarities among list words are comparable with those of the CERAD wordlist[18]; and (6) neither homophones (eg, bare/bear) nor words ending in the same phoneme (eg, plain/airplane) are allowed in a wordlist.[19]

HBCP Model for Delayed Recognition Memory

Figure 1 shows the SDT model of the memory strength distributions for *old* and *new* list words (Fig. 1A) that was incorporated into the HBCP model (Fig. 1B). Each presented word evokes a memory strength, which the subject compares with their criterion level, k, for decision making. The model predicts that a subject will respond to a word that evokes a memory strength greater than k as an *old* word, whereas a word that evokes a memory strength less than or equal to k is responded to as a *new* word. The *discriminability*, d', is the difference between mean memory strengths of old and new list words, and indicates the memory gain from studying the *old* list words. The *hit rate* is the area, h, of the *old* word distribution, and lies above k. The *false-alarm rate* is the area, f, of the new word distribution, and also lies above k. The *response bias*, c, for a subject is the distance between their criterion memory strength level, k, and the midpoint of their *discriminability*, d'. These measures of *response bias* and *discriminability* have been proposed to be independent.[3] Because recognition memory experiments have found that the SDs of the *old* and *new* word distributions differ by about 25%,[20] we incorporated this unequal variance assumption into the SDT model.

Model Extension for Group and Individual Differences

Unlike previous SDT applications to recognition memory data of ADRD patients, individual differences were modeled by introducing a parameter reflecting the 6 functional severity levels (FAST stages 1 to 6), which influenced the *response bias*, c_j, and *discriminability*, d'_j of each subject, j. Each subject's *discriminability* and *response bias* parameters were therefore drawn from the distribution of values generated by the subject's FAST stage group. In this way, the HBCP model allows different parameter values for individuals with the same FAST stage.

Model Extension for Predicting Changes in Discriminability

Discriminability, d', between *old* and *new* words was modeled by a psychophysical function that made d' a function of FAST stage severity. For FAST stage i, the mean discriminability is:

$$\mu_{d'i} = k + \left(\frac{l}{1+ae^{bi}}\right)$$

where k corresponds to baseline *discriminability*, l corresponds to the potential change in discriminability across FAST severity levels, and a and b are parameters that control the shape of the psychophysical function. A sigmoid form of the psychophysical function was selected because

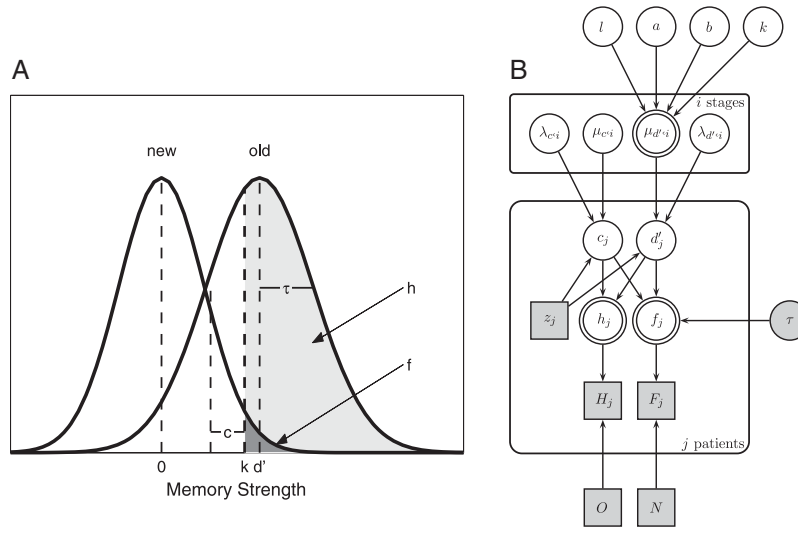

FIGURE 1. A, Signal detection theory (SDT) model. B, The hierarchical bayesian cognitive processing (HBCP) model. SDT model: the SDT model shows the memory strength distributions for *old* (studied) and *new* (unstudied) list words, along with parameters, *k* (*subject's response criterion*), *d'* (*discriminability*), *c* (*response bias*), *h* (*hit rate*), and *f* (*false-alarm rate*). The HBCP model: this HBCP model generates each patient's observed response data (*Hj*, *Fj*) from their cognitive processes of *discriminability* and *response bias*. It also models the continuum of functional severity underlying the discrete functional assessment staging test stages. See Methods section for details.

the changes in discriminability between each of the FAST stages from 1 to 6 are nonlinear (Fig. 3). The term in the denominator, ae^{bi}, can approach "0" nonlinearly as one progresses from FAST stage 1 to 6, which maximizes the value of the numerator, l, at FAST stage 6. Thus, the change in mean discriminability from baseline at FAST stage 1 will be maximal at FAST stage 6 and approach it nonlinearly. This approach goes beyond simply testing for a significant difference in *discriminability* between FAST stages and models how *discriminability* changes with functional severity.

HBCP Graphical Model Implementation

We implemented the aforementioned hierarchical SDT model in the form of a Bayesian graphical model, a formalism widely used in statistics and computer science.[21,22] In graphical models, nodes correspond to variables and their interdependencies show the causal relationships between the variables. In particular, graphical models show how unobserved variables (ie, parameters) generate observed variables (ie, data). Details and tutorials for the use of graphical models are available.[23,24] The practical advantage of graphical models is that sophisticated and general-purpose Markov chain Monte Carlo (MCMC) algorithms exist that can sample from the full joint posterior distribution of the parameters conditional on the observed data. In Bayesian statistics, the *posterior* is the probability distribution of unobserved values (often parameters) that results after a prior distribution has been updated by data.[25]

It is easiest to understand the HBCP graphical model in Figure 1 by starting with the *jth* patient's *discriminability* and *response bias* parameters (d'_j and c_j nodes). These parameters generate the *jth* patient's predicted hit, h_j, and false-alarm, f_j, rates, according to the SDT model. The hit rate is $h_j = \Phi[(d'_j/2) - c_j]$ and the false-alarm rate is $f_j = \Phi[-[(d'_j/2) + c_j]/\tau]$, where $\tau = 0.8$, arising from the unequal variance assumption. On the basis of these hit and false-alarm rates and the $O = 10$ old and $N = 10$ new words presented to each patient during the recognition task, the *jth* patient produces H_j observed hits and F_j observed false alarms, which follow binomial distributions parametrized by hit and false-alarm rates and by their number of old and new words presented [$H_j \sim \text{binomial}(h_j,O)$, $F_j \sim \text{binomial}(f_j,N)$]. Each FAST stage, i, has its own set of Gaussian distributions for the discriminability (d') and response bias (c) parameters, which are controlled by their mean, μ, and precision, λ ($\lambda = 1/\sigma^2$) variables. These FAST stage group distributions are implemented using an indicator variable, z_j, which takes the value, 1, 2,…6 according to the *jth* patient's FAST stage. For this patient, j, their *discriminability* is distributed as $d_j \sim \text{Gaussian}(\mu_{d',zj}, \lambda_{d',zj})$ and their *response bias* is distributed as $c_j \sim \text{Gaussian}(\mu_{c,zj},\lambda_{c,zj})$. Finally, the psychophysical function determines the mean discriminability of FAST stage, i ($\mu_{d'i}$ node), which then updates the subject's discriminability, d'_j.

Bayesian Inference Generated by the HBCP Graphical Model and Clinical Data

Graphical HBCP modeling was performed using WinBUGS software.[26] This software uses a range of MCMC computational methods to obtain samples from the posterior distributions of the relevant parameters.[27] To perform all analyses, 10,000 posterior samples were collected, after a burn-in of 1000 samples (samples collected but not used to approximate the posterior distribution of interest), using multiple chains to check convergence.

RESULTS

HBCP Grahical Model Fit

Posterior predictive distributions assess the descriptive adequacy of a Bayesian model by predicting what the data's distribution should be.[25] A poor-fitting model will produce inaccurate posterior predictive distributions of the data. Figure 2 shows a posterior predictive analysis for the implemented HBCP model. Rows 1 to 3 correspond to the (1) observed delayed recognition data (y axis = hits, x axis = false alarms); (2) group-level (FAST stage) model inferences; and (3) individual-level model inferences, showing 1 patient per FAST stage, with their mean value marked as an "×." The hollow black squares show the distribution of predicted hits and false alarms for each FAST stage (column). For row 1, the gray X are the patient-observed hits and false alarms. For row 2, the hollow black squares show the posterior predictive distribution of hits and false alarms at the group level (FAST stage); each square's area is proportional to its predictive mass. Comparison of rows 1 and 2 indicates that the group-level predictions match the

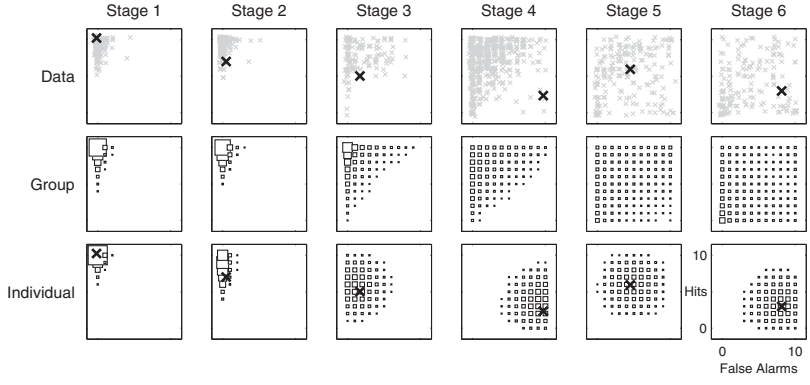

FIGURE 2. Posterior predictive assessment of the hierarchical signal detection theory model's fit to the clinical data. *Hit* and *false-alarm* distributions per functional assessment staging test (FAST) stage: observed versus group-level and individual-level posterior predictions. The predicted, individual-level data (row 3) model individuals better than group-level predictions (row 2). See Results section for details.

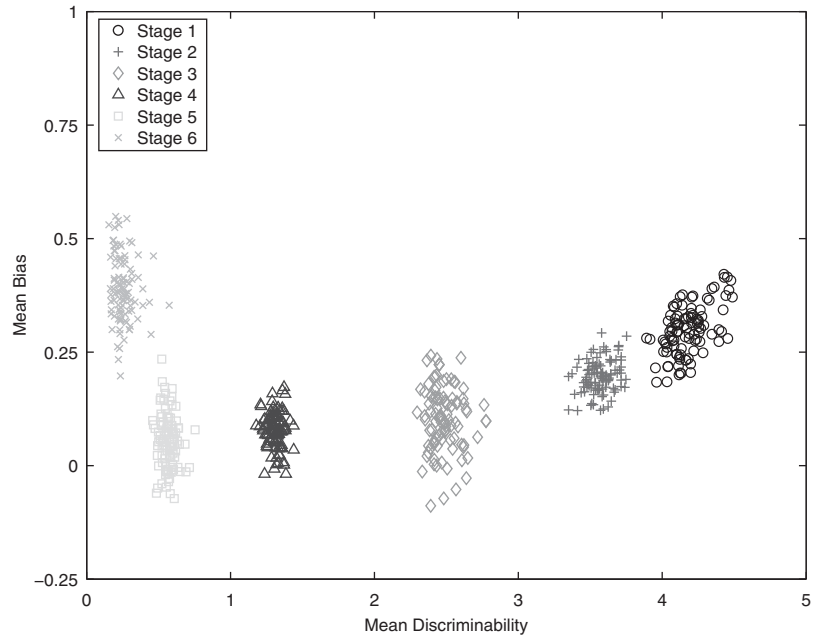

FIGURE 3. Joint posterior distributions for functional assessment staging test (FAST) stage group-level discriminability and response bias parameters. The joint posterior predictive group-level *Discriminability* and *Response Bias* parameter distributions per FAST stage. The joint posterior group-level distributions of these latent cognitive processing parameters (*discriminability* and *response bias*) completely separate the 6 FAST stages, whereas the observed behavioral data do not. See Results section for details.

observed data fairly closely, which is consistent with a well-fitting model. For row 3, the hollow black squares show the posterior predictive distribution of hits and false alarms for a selected patient in each FAST stage; each square's area is proportional to its predictive mass for that patient in that FAST stage.

Note that the posterior predictive distributions of the individuals selected for FAST stages 4 to 6 represent outliers for their FAST stage groups. The use of an individual who is an outlier for a given group illustrates the point that one can simultaneously examine both the group and the individual posterior predictive distributions. It also illustrates the point that the distribution of the individual outlier patient is different from that of the group and is more informative than simply using the group distribution for that individual. Specifically, one can see that the HBCP model's posterior predictive distribution of hits and false alarms for any given selected individual outlier patients is a much better fit than the group-level predictions in row 2. The HBCP model's ability to characterize these individuals well, while simultaneously describing group-level performance well, highlights an important advantage of the hierarchical approach for modeling individual differences.

Assessing Discriminability, Response Bias, and Changes As ADRD Progresses

Figure 3 shows the joint posterior distributions of the *discriminability* and *response bias* parameters for each FAST stage. As the degree of functional impairment increases from FAST stages 1 (circles) to 5 (squares), *discriminability* between *old* and *new* list words decreases, and *response bias* shifts toward *misses* and *false alarms* being equally likely. However, *response bias* during FAST stage 6 (Xs) shifts back toward that seen in normal aging patients (FAST 1), who make more *misses* than *false alarms*.

Figure 4 shows the fit of the modeled relation between *discriminability* and the degree of impairment in functional capabilities (FAST stage). MCMC sampling of the *discriminability* values, d'_i, generated by the psychophysical function, was used to estimate the mean discriminability (black curve) and its 95% credible interval (thick blue lines) per FAST stage. The uncertainty of predicted *discriminability* was also estimated by random sampling of the d' parameter values from their associated posterior distributions (gray curves). Both measures of uncertainty showed that the predicted *discriminability* of each FAST stage was reliably predicted by the continuous measure of the severity of functional impairment modeled by the psychophysical function.

DISCUSSION

The HBCP model of delayed recognition memory showed that decision making—an executive function modeled by *response bias*—shifts toward unbiased responding during subjective cognitively related functional impairment and MCI (FAST 2 to 3). A surprising result was the shift back to biased responding similar to NCI individuals (FAST 1) during moderately severe dementia (FAST 6). FAST stage 6 patients have severe memory impairment, so that the memory strength distributions for the *old* and *new* wordlists will be similar. This loss of discriminability means that there is no *memory* signal to make a decision between old and new list words. However, *judgment* becomes more severely impaired as dementia severity progresses from FAST stages 4 to 6. This decline in judgment may shift the FAST stage 6 patient's decision-making

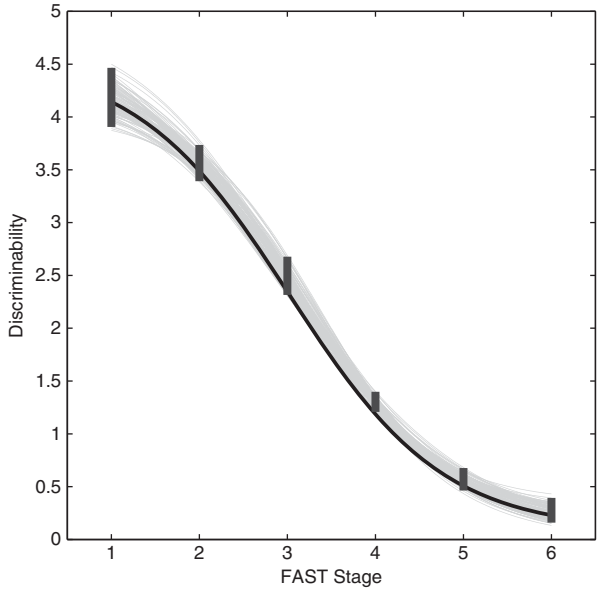

FIGURE 4. Predicted relation between discriminability and functional assessment staging test stage transformed into a continuous measure of functional severity. Modeling the underlying continuum of functional impairment. The psychophysical function that uses the *discriminability* cognitive processing parameter to model the continuum of functional impairment underlying the discrete functional assessment staging test stage values fits well. See Results section for details.

criterion, k, to the right. This shift would yield a response bias, c, similar to that seen in NCI (FAST 1).

A potentially useful clinical application is that delayed recognition memory tasks can be used to create a continuous measure of severity of functional impairment that reliably predicts FAST staging, which is an ordinal measure. A continuous measure of functional impairment allows one to compute the rate of functional decline, which can be used, for example, to determine whether a treatment has delayed disease progression.

The relatively good fit between discriminability, d', and the FAST stages, as shown in Figure 4, means that the psychophysical function used to model this relationship will aid the interpolation, generalization, and prediction of the severity of functional impairment. In other words, this psychophysical function allows one to trace out trajectories of functional decline with respect to discriminability and map these trajectories into statements about memory task performance.

The HBCP model presented here demonstrates how one can simultaneously evaluate clinically relevant groups (ie, FAST stage groups) and individuals within each group. The individual-level fits (see row 3, Fig. 2) show that one can predict the distribution of an individual's recognition memory performance better than that obtained by the individual's group-level predictions. This is particularly useful for patients who may belong to a distinct subset of the distribution.

The HBCP model presented here also shows how latent processes of memory and executive function that are not directly measurable can be usefully estimated from the delayed recognition memory response data. These latent parameters separated the group-level values for FAST stages 1 and 2 (Fig. 3), whereas the observed recognition memory data did not (see row 1, Fig. 2). This improved separation of the FAST stages illustrates an important advantage of generative HBCP models over discriminative statistical methods.

HBCP models accurately translated a latent measure of cognition into a continuous measure of the degree of impairment in functional capabilities. This translational ability may facilitate better understanding of the relations between cognition, function, and other levels of brain information processing, including those measured by biomarkers at molecular, structural, and electrophysiological levels. In this regard, our future studies will examine how this continuous measure of functional severity relates to ADRD diagnosis, quantitative MRI volumetric data, apolipoprotein E genotype, cerebrospinal fluid levels of Abeta42, phospho tau, and total tau, plus affective states such as depression, and behavioral states such as agitation, aggression, and psychosis.

REFERENCES

1. MacMillan N, Creelman CD. *Detection Theory: A User's Guide*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates Inc.; 2005.
2. Budson AE, Wolk DA, Chong H, et al. Episodic memory deficits in Alzheimer's disease: separating response bias from discrimination. *Neuropsychologia*. 2006;44:2222–2232.
3. Snodgrass JG, Corwin J. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol Gen*. 1988;117:34–50.
4. Sclan SG, Reisberg B. Functional Assessment Staging (FAST) in Alzheimer's disease: reliability, validity and ordinality. *Int Psychogeriatr*. 1992;4(suppl 1):55–69.
5. Reisberg B, Jamil IA, Khan S, et al. Staging dementia. In: Abou-Saleh MT, Katona C, Kumar A, eds. *Principles and Practice of Geriatric Psychiatry*. 3rd ed. Chichester, England: John Wiley & Sons Ltd.; 2011:162–169.
6. Reisberg B. Functional assessment staging (FAST). *Psychopharmacol Bull*. 1988;24:653–659.
7. Reisberg B. Dementia: a systematic approach to identifying reversible causes. *Geriatrics*. 1986;41(4):30–46.
8. Shankle WR, Romney AK, Hara J, et al. Method to improve the detection of mild cognitive impairment. *Proc Natl Acad Sci USA*. 2005;102:4919–4924.
9. Trenkle D, Shankle WR, Azen SP. Detecting cognitive impairment in primary care: performance assessment of three screening instruments. *J Alzheimers Dis*. 2007;11:323–335.
10. Cho A, Sugimura M, Nakano S, et al. The Japanese MCI Screen for early detection of Alzheimer's disease and related disorders. *Am J Alzheimers Dis Other Demen*. 2008;23:162–166.
11. Shankle WR, Mangrola T, Chan T, et al. The CERAD wordlist memory performance index: development and validation. *Alzheimers Dement*. 2009;5:295–306.
12. Brun A, Passant U. Frontal lobe degeneration of non-Alzheimer type. Structural characteristics, diagnostic criteria and relation to other frontotemporal dementias. *Acta Neurol Scand Suppl*. 1996;168:28–30.
13. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol*. 2007;6:734–746.
14. Walker Z, Jaros E, Walker RW, et al. Dementia with Lewy bodies: a comparison of clinical diagnosis, FP-CIT single photon emission computed tomography imaging and autopsy. *J Neurol Neurosurg Psychiatry*. 2007;78:1176–1181.
15. Wiederkehr S, Simard M, Fortin C, et al. Validity of the clinical diagnostic criteria for vascular dementia: a critical review. Part II. *J Neuropsychiatry Clin Neurosci*. 2008;20:162–177.

16. Mayeux R, Reitz C, Brickman AM, et al. Operationalizing diagnostic criteria for Alzheimer's disease and other age-related cognitive impairment-Part 1. *Alzheimers Dement*. 2011; 7:15–34.
17. Nelson DL, McEvoy CL, Schreiber TA. The University of South Florida free association, rhyme, and word fragment norms; 1998. Available at: http://web.usf.edu/FreeAssociation/Intro.html. Accessed date, February 20, 2012.
18. Landauer TK, Dumais ST. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychol Rev*. 1997;104:211–240.
19. Weide R. The Carnegie Mellon Pronouncing Dictionary v0.4. Carnegie Mellon University. 1996. Available at: ftp://ftp.cs.cmu.edu/project/fgdata/dict. Accessed date, February 20, 2012.
20. Mickes L, Wixted JT, Wais PE. A direct test of the unequal-variance signal-detection model of recognition memory. *Psychon Bull Rev*. 2007;14:858–865.
21. Jordan MI. Graphical models. *Stat Sci*. 2004;19:140–155.
22. Koller D, Friedman N, Getoor L, et al. Graphical models in a nutshell. In: Getoor L, Taskar B, eds. *Introduction to Statistical Relational Learning*. Cambridge, MA: MIT Press; 2007:13–56.
23. Lee MD. Three case studies in the Bayesian analysis of cognitive models. *Psychon Bull Rev*. 2008;15:1–15.
24. Shiffrin RM, Lee MD, Kim WJ, et al. A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cog Sci*. 2008;32:1248–1284.
25. Gelman A, Carlin JB, Stern HS, et al. *Bayesian Data Analysis*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC; 2003:3–32.
26. Spiegelhalter DJ, Thomas A, Best NG. *WinBUGS Version 1.4 User Manual*. Cambridge, UK: Medical Research Council Biostatistics Unit; 2004.
27. MacKay DJC. *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press; 2003.