Published in final edited form as:

Alzheimer Dis Assoc Disord. 2009 ; 23(1): 44–49. doi:10.1097/WAD.0b013e31817634a0.

THE “PORTABLE” CDR: TRANSLATING THE CLINICAL DEMENTIA RATING INTERVIEW INTO A PDA FORMAT

James E. Galvin, MD, MPH¹, Thomas M. Meuser, PhD¹, Mary A. Coats, MSN¹, Donald A. Bakal, PhD^{2,3}, and John C. Morris, MD¹

¹ Alzheimer's Disease Research Center, Department of Neurology, Washington University School of Medicine, St. Louis

² Rockyview General Hospital, Calgary, Alberta, Canada

³ Department of Medicine, University of Calgary, Alberta, Canada

Abstract

The Clinical Dementia Rating (CDR) is a common rating system used in clinical trials and longitudinal research projects to rate the presence and severity of cognitive problems in Alzheimer disease and related disorders. The interview process requires training and can be time consuming. Here we describe the validity, reliability and discriminative ability of a computer-generated CDR using a personal digital assistant (PDA) format. This project utilized clinical data from 138 archival and live evaluations (patient and informant interviews) collected for research purposes at Washington University to develop and test a software-based system for the administration and automatic scoring of the CDR. The system was programmed for use on a hand-held computer via the Palm™ Operating System. We developed domain-specific algorithms to quantify and translate clinical scoring decisions for the three cognitive (Memory, Orientation, Judgment and Problem Solving) and the three functional (Community Affairs, Home & Hobbies, Personal Care) domains of the CDR. An acceptable set of algorithms were developed using data from 104 research cases, reflecting a range of impairment levels (CDR 0 – 3) and expert scoring decisions. These algorithms were then tested for accuracy in a validation sample of 34 cases. The computer-generated CDR has excellent internal consistency (Cronbach's alpha ranging from .94–.98) and inter-rater reliability (intraclass correlation coefficient ranging from .88 – .96). The computer-generated CDR showed excellent discrimination between demented and nondemented cases (Area under the curve = 0.95; 95% CI: .84–1.1). The computer-generated CDR using a Palm™ Operating System is easy to use, valid and reliable. The level of agreement compares favorably to published inter-rater reliability data for the CDR. Software-based administration and automatic scoring of the CDR is a viable alternative to paper-based methods and may be useful in research and clinical settings, especially where electronic data management and reliability in scoring are critical.

Keywords

Clinical Dementia Rating; CDR; Alzheimer's; Dementia; Impairment; Staging

Address correspondence and reprint requests to: James E. Galvin, MD, MPH, Alzheimer Disease Research Center, Washington University School of Medicine, 4488 Forest Park, Suite 130, St. Louis, Missouri 63108, Telephone: (314) 286-2547, Fax: (314) 286-2763, galvinj@neuro.wustl.edu.

DISCLOSURE

Scoring algorithm and software development was performed with assistance from Minding Digits, Inc (<http://www.mindingdigits.com/>). Janice Bakal, the Founder and President of Minding Digits, Inc., is the spouse of co-author Donald A. Bakal. None of the authors received financial support from Minding Digits, Inc.

INTRODUCTION

The Clinical Dementia Rating (CDR) is a commonly used clinical staging instrument characterizing the presence and severity of dementia in Alzheimer disease (AD) and related disorders.^{1,2} The CDR is generated from a semi-structured interview with the patient and a knowledgeable collateral source (CS), such as the spouse or adult child. The CDR is derived using information from the clinical assessment but without reference to psychometric performance, and rates cognitive function in six categories (Memory, Orientation, Judgment and Problem Solving, Community Affairs, Home and Hobbies and Personal Care). The global CDR is derived by synthesizing ratings in each of the six categories where CDR = 0 indicates no dementia, CDR = 0.5 signifies uncertain or very mild dementia^{3,4}, and CDR = 1, 2, or 3 corresponds to mild, moderate, or severe dementia.^{1,2} The CDR scoring table includes descriptive “anchors” that guide the rater in determining individualized domain ratings based on interview data and clinical judgment. Domain scores are intended to reflect cognitive, behavioral and functional changes resulting from dementia. When administered in a standardized manner by a trained rater, the CDR is a valid and reliable tool for staging dementia-related impairment.^{5–8}

Unlike norm-based psychometric measures which compare individual functioning to a reference group, the CDR measures intra-individual change from pre-dementia baseline. Dementia presentation can vary widely across individuals, thus requiring an individualized approach. An individualized approach negates confounding factors (age, education level, race, etc.) and ceiling/floor effects that limit many norm-based measures.⁹ The CDR is reliant on informant data. Persons with dementia often lack insight into their deficits, making informant data quite helpful for confirming true cognitive-functional changes from past baseline.

There are several disadvantages to current methods of generating the CDR. The administration and scoring of the CDR is conducted through live interviews with patient and informant; data is recorded on paper and scored. While this format may provide information to support clinical judgment, the CDR interview can be time-consuming (taking 45–60 minutes) and therefore impractical for many clinical and research settings. In addition, paper-based data can be problematic when computer analysis or electronic charting is necessary. Transference of paper records into a computerized database adds time and the potential for error to the process.

To improve the efficiency of the CDR in its administration, scoring and data management, and thus improve its utility, a collaborative project was initiated to develop, test and validate a computer software-based system. Such a system needed to be built on a readily available platform; user friendly and capable of being operated correctly with minimal training; fully automated yet also modifiable in response to clinical judgment; and capable of electronic data transfer for desktop computer storage and analysis. Of the range of available devices, the Palm Pilot™ platform and associated operating system (OS) were chosen for programming of the prototype software. These hand-held computer devices have been available in the marketplace for over a decade, and have good connectivity to the desktop PC. In addition, many software titles exist for the Palm OS, including programs that connect to Microsoft data management products, such as Excel and Access.

METHODS

Sample Description

Project team members translated the CDR interview into a software application for hand-held computer administration. The computerized interview includes user-friendly entry screens for patient demographics, interview questions, and content notes. The computer algorithm for obtaining the global score is also incorporated. Four independent samples were used in this

project. The CDR interviews were entered by hand by researchers while watching the interview on a TV screen to simulate how the Palm would operate when clinical information was entered as part of a real interview. The scoring algorithm was first developed (Figure 1) using 9 training videos from Washington University's Brief Training & Reliability Protocol (BTRP) for the CDR conducted by a single CDR expert (JCM). (<http://alzheimer.wustl.edu/cdr/>). The nine BTRP interviews include examples of both normal cognitive aging and AD, and cover all CDR impairment levels (two interviews for CDR = 0, 0.5, 1 and 2; one interview for CDR = 3).

To test the generalizability of the initial domain algorithms, an independent set of 95 cases were randomly selected from the Washington University research files. This new sample of cases reflected all impairment levels (CDR 0 = 24; CDR 0.5 = 24; CDR 1 = 22; CDR 2 = 15; CDR 3 = 10). In addition, a variety of dementia etiologies were studied (24 nondemented; 53 AD; 8 uncertain dementia; 10 other disorders (including Parkinson Disease, Vascular Dementia, Dementia with Lewy Bodies). In this sample, interviews were conducted and scored by different clinicians (neurologists, geriatricians, psychiatrists and masters-prepared nurse clinicians).

Lastly, 2 validation samples were used. Data from 20 cases from the Washington University research files were randomly selected reflecting a variety of CDR stages (CDR 0 = 3; CDR 0.5 = 9; CDR 1 = 4; CDR 2 = 2; CDR 3 = 2) and diagnoses (3 nondemented, 16 AD, 1 uncertain dementia). In addition, a trained nurse clinician utilized the system for 14 live data collection during research evaluations in the community. The following impairment levels were represented: CDR 0 = 7; CDR 0.5 = 3; CDR 1 = 1; CDR 2 = 1; CDR 3 = 2. Primary diagnoses included 7 nondemented, 2 uncertain dementia and 5 AD individuals. Entries were made to the computer while an independent and blinded clinician administered the CDR in the standard fashion on paper. The clinician entering this live data did not alter scores based on clinical judgment, allowing the computer-based algorithms to make all determinations.

Algorithm Development Sample

The approach used to develop and refine domain-specific algorithms did not lend itself to traditional statistical methods, such as logistic regression. Computer logic needed to reflect the "binary" thinking and decision rules built into the CDR. The approach adopted resembles the logic of Classification and Regression Tree (CART) analysis with the CDR domain scores viewed as categorical outcomes and the predictor variables being item responses to corresponding interview questions. For example, in the CDR domain "Judgment and Problem Solving", the informant is asked to report on the observed functioning of the patient (e.g., "Can he/she handle a household emergency – plumbing leak, small fire?") and the patient is tested with questions on similarities/differences between objects, mathematical calculation, and common sense judgment (e.g., "Upon arriving in a strange city, how would a person locate a friend there that they wished to see?"). Response patterns to these questions were examined across many cases to identify a set of rules that would accurately predict assigned clinical scores (Table 1). As model cases for training, the BTRP patient/informant interviews were useful for developing initial scoring algorithms for each domain. Decision patterns linking interview content to selected clinical scores could be readily discerned and applied in rule form.

Scoring rules from the initial algorithms did not initially translate well to "real world" cases. Agreement rates between computer and clinician were below 50%. Conventional CDR scoring requires that all information be used to create single scores for Memory, Orientation and Judgment and Problem-solving, but these are based on data from two sources: patient and informant. It soon became clear that clinician interviewers in the 95 case sample assigned different weight or emphasis to information obtained from informant and patient when generating specific domain scores. In response to this observation, the computer algorithm was adapted to generate separate patient and informant scores for three cognitive domains

(Memory, Orientation, Judgment and Problem-solving) in order to maintain the integrity of the original decision-making process. A single informant score was used for the remaining three functional domains (Community Affairs, Home and Hobbies, Personal Care). With this method, differential weighting could be included, such that patient and informant domain scores, once calculated, may be more accurately combined to obtain overall domain scores.

Table 1 illustrates the first step in this process for the *Judgment and Problem Solving* domain with regards to interview data from the patient. Three separate item clusters were created based on this patient content: (1) similarities and differences (S & D), (2) calculation, and (3) judgment. The classification criteria were determined so as to most accurately predict the “gold standard” scores of Washington University raters. A similar process was followed to derive the informant domain score for *Judgment & Problem Solving* and the other domain scores (Memory, Orientation, Home & Hobbies, Community Affairs, Personal Care).

Generalizability Sample

In the next step, the team developed an additional algorithm to merge patient and informant domain scores. A simple average was unreasonable, as half and quarter points do not exist in the CDR. Instead, specific scoring rules were assigned based on patterns of clinical judgments present in the 95 case sample. A tangible strength of the CDR interview lies in its ability to collect informant data and apply this to understanding dementia status. Persons with dementia typically lack insight into their deficits, making some patient-derived information suspect. In the experience of Washington University clinicians, clinical judgment will often favor informant data in making scoring decisions.¹⁰ In general, the algorithm for each final domain score reflects this reality by assigning more weight to informant data at the low end of the severity continuum (0, 0.5, 1); and, conversely, more weight to patient data at the higher end of impairment to account for multiple errors (2, 3). The algorithm to determine final domain score for *Judgment and Problem Solving* is depicted in Table 2.

Statistical Analysis

All analyses were performed using SPSS v.13.0 (Chicago, IL). Validity was assessed comparing the performance of the Palm algorithms with a “gold standard” – in this case, the BTRP using Spearman correlation coefficients.⁹ Receiver operator characteristic (ROC) curves and the area under the ROC curve (AUC) were generated to reflect graphically and quantitatively the ability of the Palm algorithms to discriminate between nondemented patients (CDR 0) and patients with dementia (CDR ≥ 0.5). The sensitivity, specificity, positive and negative predictive values of the Palm algorithms were calculated.

Internal consistency was examined as the proportion of the variability in the responses that is the result of differences in the rating format (clinician vs. Palm). Internal consistency was reported as Cronbach’s alpha reliability coefficient.⁹⁻¹¹ Intraclass correlation coefficients (ICC) were calculated to assess the percent agreement between raters (inter-rater reliability). The ICC is computed from multiple observations of the same variable to evaluate the consistency among the raters rather than absolute agreement.⁹⁻¹¹ Simple agreement (i.e. the proportion of responses in which two observations agree) is strongly influenced by the distribution of positive and negative responses, as well as the possibility of agreement by chance alone.¹¹ We used the scheme reported by Fleiss in assessing agreement.¹² An ICC between 0.55 and 0.75 was considered good agreement while an ICC statistic greater than 0.76 was considered excellent.¹²

RESULTS

Algorithm Development

A preliminary set of domain algorithms were programmed based on BTRP interview material, achieving an absolute agreement rate of 89% (8 of 9 correct) comparing the global CDR scores between computer and BTRP “gold standard” ratings. Concurrent (Criterion) validity is the correlation of the computer-generated CDR compared with “gold standard” CDR from the BTRP to assess the ability of the computer-generated CDR to adequately differentiate groups that should be differentiable.^{9,11} The computer-generated CDR was highly correlated with the BTRP CDR (Spearman correlation 0.974, $p < .001$) For the one case in error, the computer rated the patient’s cognitive-functional impairment as moderate (CDR = 2), whereas the “gold standard” rating was in the mild range (CDR = 1).

Generalizability Sample

The combined domain-specific algorithms produced a good match with the clinician generated scores, ranging from 77 – 93% (Table 3). Also, the combination algorithms resulted in more accurate scores (relative to clinician determined scores) than informant or patient domain scores alone. With regards to the Global CDR Score, the software-generated domain scores yielded correct global scores via the published Washington University algorithm in 86% (82 of 95) of the cases.

We tested the degree to which the Computer-generated CDR scores were free from random error by assessing the internal consistency with Cronbach’s alpha.^{9,11} This provides an estimate of reliability based on all possible correlations between items. Cronbach’s alpha for the individual domains was good, ranging from 0.86 for the Memory domain to 0.66 for Orientation (Table 4). Inter-rater reliability was assessed with ICC statistics (Table 3). There was good agreement between the clinician and computer-generated CDR scores in every domain except Orientation.

Team members then met to review the 13 (of 95) cases that did not match, focusing on rule modifications to enhance accuracy. In most instances, however, data outside of the structured CDR interview questions (i.e. open-ended questions) was considered primarily responsible for scoring decisions, suggesting that the clinician had the benefit of live interaction with the patient and information that influenced his/her scoring decisions. Such experience cannot be programmed. The core set of decision rules could not be altered without creating scoring deviations in the sample, however an option for a manual override option was incorporated to allow for clinician judgment to be integrated into the final CDR score.

Validation Samples

To validate the computer-generated CDR, data from 20 recent archived research cases and 14 “real-time” CDR interviews were evaluated. Results of these validation efforts are presented in Table 4. The percentages of agreement between clinician and software-generated domain scores range from 68–85% with agreement higher in the live administration cases. Among the four cases that did not match, the software underestimated impairment by one level in two cases (assigning CDR 1 to a CDR 2 case, and a CDR 2 to a CDR 3 case), and overestimated in two cases (assigning a CDR 0.5 to a 0 case, and a CDR 1 to a 0.5 case). The agreement in global CDR between raters was 88%, corresponding favorably to published inter-rater reliability showing 83% agreement in global CDR rating.⁵

Internal consistency in the validation samples was excellent with Cronbach’s alpha ranging from .94–.98 (Table 4). Inter-rater reliability was also excellent for individual CDR domains ranging from .88 for Judgment and Problem-solving to .95 for Orientation. The ICC for the

Global CDR was 0.96 (95%CI: .92–.98) suggesting excellent agreement between the clinician and computer-generated CDR.

Discriminative Properties of the Computer-generated CDR in the Validation sample

ROC curves were generated to measure the ability of the computer-generated CDR to correctly classify CDR 0 (nondemented) vs. CDR ≥ 0.5 (demented) in the validation samples ($n=34$). For the comparison of CDR 0 vs. CDR ≥ 0.5 , the area under the curve was 0.95 (95% CI: 0.84–1.06, $p<.001$), suggesting excellent ability to discriminate between nondemented and demented individuals. When only CDR 0 and CDR 0.5 cases ($n=23$) were considered, the AUC = 0.95 (95% CI: .84–1.1, $p<.001$) suggesting that the Palm-based CDR had excellent ability to discriminate very mild cases of cognitive impairment often presenting the greatest clinical challenge. The software prototype showed a sensitivity of 100%, and a specificity of 96%. With a dementia prevalence of 47%, the positive predictive value was 90%, and the negative predictive value was 100%.

DISCUSSION

This project utilized clinical data from archival and live evaluations (patient and informant interviews collected for research purposes; $n = 138$) to develop, test and validate a software-based system for the administration and automatic scoring of the Clinical Dementia Rating (CDR). The system was programmed for use on a hand-held computer via the Palm Operating System (OS) – a readily available platform with good connectivity for data transfer to desktop applications. After programming CDR items into a user-friendly format for data entry, a primary challenge was the development of domain-specific algorithms to accurately translate interview data for three cognitive (Memory, Orientation, Judgment and Problem Solving) and three functional (Community Affairs, Home & Hobbies, Personal Care) domains of the CDR.

After defining a set of acceptable algorithms, we piloted use of the computer-generated CDR in the field and demonstrated excellent validity and reliability properties. The computer-generated CDR was highly correlated with the traditional clinician-generated CDR and had high internal consistency and inter-rater reliability. The computer-generated CDR also showed excellent discrimination between demented and non-demented cases compared with the clinician-generated CDR. This level of agreement compares favorably to published inter-rater reliability data for the CDR and suggests that computer-based administration and scoring may be a viable option in many settings.⁶ Of note, the prototype software system was especially accurate for distinguishing response patterns between CDR 0 and 0.5 cases, the most challenging distinction for practicing clinicians.

The strengths of the computer-generated CDR included its accuracy and ease of use. In all the cases tested in developing and testing this prototype system, no error in global CDR score was greater than one level. In other words, when the prototype was incorrect, it missed by a close margin. An individual with mild impairment may be labeled as very mildly impaired or moderately impaired, but not severely impaired. Of note, during development and validity testing, only 1 nondemented CDR 0 case was assigned a CDR 0.5 cases and no cognitively impaired individuals were rated as CDR 0, nondemented by the Palm software. Infrequent errors of this type may be acceptable in some settings, especially when other sources of information are available to guide clinical judgment. In the area of clinical evaluation and care, it is common to collect more information about a patient and his/her current status than just covered by the CDR. The generally accurate rating provided by the prototype software system may prove quite helpful in such settings. Before an investigator scores the CDR in the “traditional” fashion, the rater is generally required to undergo a period of training, the same BTRP that was used to score the Palm CDR. The Palm platform provides structure so that

research assistants and non-physician clinicians should be able to use it with little additional training (mainly familiarity with the platform and data entry).

There are limitations to this study. Despite these promising findings, further validation is needed to confirm the accuracy of this prototype software. A larger validation sample must be examined, including more non-AD dementia cases and with clinician raters from a broader spectrum of settings and professional backgrounds. Although the prototype software system for the CDR appears to be accurate, the evidence for this to date is based on data collected by experts in the traditional means of CDR administration and scoring. The global rating from this automated system will only be as accurate as the skills of the interviewer will allow. If the interviewer or rater has limited knowledge of dementia, or is unable to probe effectively when additional information is needed, what is entered to the software could be incorrect and, thus, could adversely influence the calculation of the global score. However, the traditional CDR scoring typically requires training. This training permits standardization of inter-rater reliability.¹³ Similar training for the computer-algorithm CDR would probably be required.

It is also unclear how clinical interviewers and raters may respond to software-generated domain scores and what use clinicians may make of the option to over-ride the computer to incorporate clinician judgment. Previous research has supported that informants are reliable in detecting the earliest signs of dementia.¹⁰ The algorithm does take into account a rating of the informant's reliability and asks the clinician to adjust the CDR accordingly. The algorithm was designed to improve the ease of administration and reliability of scoring, not to "replace" clinical judgment.

For the research community, a tangible benefit of this system lies in its efficiency of data collection and transfer. It is not necessary, for example, for the rater to be a physician or other highly trained professional. So long as the rater has good interviewing skills and knows how to operate the software, the information entered should be sufficient to yield an accurate global rating of impairment. The prototype software developed for the Palm OS is transferable to other platforms/devices (e.g., Laptop or Tablet PC), making this system even more attractive for use in both research and clinical settings.

We believe that the current software system represents an important advance for the evaluation of dementia-related impairment, suitable for use in a variety of community research settings. The domain-specific algorithms embedded in the prototype accurately operationalize the clinical decisions and interpretations of recognized CDR experts across a wide range of patient-informant interviews. We believe that the high level of agreement achieved thus far warrants "beta" testing of this prototype, particularly in clinical research settings where reliability of data is critical.

Acknowledgments

This study was supported by grants from the National Institute on Aging (P01 AG03991 and P50 AG05681) and the Alberta Health & Wellness. The authors wish to thank Washington University nurse clinicians for their expert feedback concerning the functionality of the CDR software and assistance in live testing. The CDR is a copyrighted instrument of the Alzheimer's Disease Research Center (ADRC), Washington University School of Medicine (<http://alzheimer.wustl.edu/>).

References

1. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *British Journal of Psychiatry* 1982;140:566–572. [PubMed: 7104545]
2. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* 1993;43:2412–2414. [PubMed: 8232972]

3. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B. Current concepts in mild cognitive impairment. *Archives of Neurology* 2001;58:1985–1992. [PubMed: 11735772]
4. Morris JC, Storandt M, Miller JP, et al. Mild cognitive impairment represents early-stage Alzheimer disease. *Arch Neurol* 2001;58:397–405. [PubMed: 11255443]
5. Rockwood K, Strang D, MacKnight C, Downer R, Morris JC. Interrater reliability of the Clinical Dementia Rating in a multicenter trial. *J Am Geriatr Soc* 2000;48:558–559. [PubMed: 10811551]
6. Morris JC, Ernesto C, Schafer K, Morris JC, Ernesto C, Schafer K, Coats M, Leon S, Sano M, Thal LJ, Woodbury P. Clinical Dementia Rating training and reliability in multicenter studies: The Alzheimer's Disease Cooperative Study experience. *Neurology* 1997;48:1508–1510. [PubMed: 9191756]
7. McCulla MM, Coats M, Van Fleet N, Duchek J, Grant E, Morris JC. Reliability of clinical nurse specialists in the staging of dementia. *Arch Neurol* 1989;46:1210–1211. [PubMed: 2818255]
8. Burke WJ, Miller JP, Rubin EH, Morris JC, Coben LA, Duchek J, Wittels IG, Berg L. Reliability of the Washington University Clinical Dementia Rating. *Arch Neurol* 1998;45:31–32. [PubMed: 3337672]
9. Galvin JE, Roe CM, Xiong C, Morris JC. The validity and reliability of the AD8 informant interview for dementia. *Neurology* 2006;67:1942–1948. [PubMed: 17159098]
10. Cacchione PZ, Powlishta KK, Grant EA, Buckles VD, Morris JC. Accuracy of collateral source reports in very mild to mild dementia of the Alzheimer type. *J Am Geriatr Soc* 2003;51:819–823. [PubMed: 12757569]
11. Streiner, DL.; Norman, GR. *Health measurement scale: A practical guide to their development and use*. Vol. 3. Oxford, England: Oxford University Press; 2003.
12. Fleiss JL, Chilton NW. The measurement of inter-examiner agreement on periodontal disease. *J Periodontal Res* 1983;18:601–6. [PubMed: 6230433]
13. Rockwood K, Strang D, MacKnight C, Downer R, Morris JC. Interrater reliability of the Clinical Dementia Rating in a multicenter trial. *J Am Geriatr Soc* 2000;48:558–559. [PubMed: 10811551]

Figure 1.
Schematic of Algorithm Development and Validation

Table 1

Patient Scores for the *Judgment and Problem Solving* Domain

CDR Level	Judgment & Problem Solving Domain	Derivation Rules
No Impairment 0	Solves everyday problems & handles business & financial affairs well; judgment good in relation to past performance	One incorrect response allowed (Total score ≤ 1)
Very Mild Dementia 0.5	Slight impairment in solving problems, similarities, and differences	At least one correct response to <i>similarities and differences</i> ($PJ1 + PJ2 + PJ3 + PJ4 < 7$) and some impairment in <i>calculation</i> ($PJ5 + PJ6 + PJ7 < 3$) and/or <i>judgment</i> ($PJ8 + PJ9 < 2$)
Mild 1	Moderate difficulty in handling problems, similarity and difficulties; social judgment usually maintained	At least one correct response to <i>similarities and differences</i> ($PJ1 + PJ2 + PJ3 + PJ4 < 7$), and severe impairment in either <i>judgment</i> ($PJ8 + PJ9 \geq 2$) or <i>calculation</i> ($PJ5 + PJ6 + PJ7 \geq 2$)
Moderate 2	Severely impaired in handling problems, similarities and differences; social judgment usually impaired	At least one correct response to <i>similarities and differences</i> ($PJ1 + PJ2 + PJ3 + PJ4 < 7$), and severe difficulties with <i>calculation</i> ($PJ5 + PJ6 + PJ7 = 3$), and <i>judgment</i> ($PJ8 + PJ9 \geq 2$)
Severe 3	Unable to make judgments or solve problems	No correct responses on <i>similarities and differences</i> ($PJ1 + PJ2 + PJ3 + PJ4 = 7$)

Key: PJ1 = Patient Judgment, Interview item 1

Table 2

Derivation of the Final Domain Score for *Judgment and Problem Solving*

CDR Level	Judgment & Problem Solving Domain	Criteria Applied
No Impairment 0	Solves everyday problems & handles business & financial affairs well; judgment good in relation to past performance	Informant = 0 or Informant = 0.5 and Patient = 0
Very Mild Dementia 0.5	Slight impairment in solving problems, similarities, and differences	Informant = 0.5 and Patient > 0
Mild 1	Moderate difficulty in handling problems, similarity and difficulties; social judgment usually maintained	Informant = 1 and Patient < 2
Moderate 2	Severely impaired in handling problems, similarities and differences; social judgment usually impaired	Informant = 1 and Patient = 2 or Informant = 2 and Patient <= 2
Severe 3	Unable to make judgments or solve problems	Informant = 3 or Patient = 3

Table 3
Reliability Statistics Comparing Clinician and Computer-generated CDR in Generalizability Sample (n=95)

CDR Domains	Percent Absolute Agreement			Cronbach's alpha	ICC (95%CI)
	Informant Score	Patient Score	Final Computed Score		
Memory	71%	69%	86%	.858	.751 (.65-.83)
Orientation	76%	73%	78%	.657	.489 (.32-.63)
Judgment	73%	66%	80%	.760	.613 (.47-.72)
Community Affairs	77%	N/A	77%	.785	.647 (.52-.75)
Home & Hobbies	93%	N/A	93%	.734	.579 (.43-.70)
Personal Care	92%	N/A	92%	.742	.590 (.44-.75)
Global CDR Score	N/A	N/A	86%	.712	.553 (.40-.68)

Key: CDR = Clinical Dementia Rating; ICC = Intraclass correlation coefficient; N/A = not applicable

Reliability Statistics Comparing Clinician and Computer-generated CDR in Validity Samples (n=34)

Table 4

CDR Domains	Percent Absolute Agreement	Cronbach's alpha	ICC (95% CI)
Memory	68%	.946	.897 (.80-.95)
Orientation	85%	.972	.946 (.89-.92)
Judgment & Problem Solving	65%	.937	.881 (.78-.94)
Community Affairs	71%	.942	.891 (.79-.94)
Home & Hobbies	82%	.950	.905 (.82-.95)
Personal Care	82%	.955	.913 (.83-.96)
Global CDR Score	88%	.978	.957 (.92-.98)