Technical Documentation for

Dementia Severity Measure

Released August 2023

Niloofar Fouladi-Nashta, Shengjia Xu, Julie Zissimopoulos

# Introduction

The goal of this project is to create a measure of dementia severity for respondents in the Health and Retirement Study (HRS) and make the measure available to all researchers. The measure is based on data from the Aging, Demographics, and Memory Study (ADAMS) and the Health and Retirement Study (HRS). This package describes how the severity measure was developed and provides programming code for replicability.

# Data Sources

The main data sources for the severity measure are the HRS and ADAMS. HRS is a longitudinal study of US adults conducted by the Institute for Social Research at the University of Michigan; The ADAMS study is a supplement to the Health and Retirement Study (HRS), which is sponsored by the National Institute of Aging (grant number NIA U01AG009740). It was conducted jointly by Duke University and the University of Michigan.

The ADAMS sample is a sub-sample of HRS participants (n=1,770) who were selected from the 2000 and 2002 waves of HRS to undergo a detailed clinical assessment of their cognitive function based on their self- or proxy-cognitive assessment measure 1. Among this group, 852 respondents (wave A) were assessed as part of the initial visit between August 2001 and December 20031. The severity measure is predicted using the 852 respondents in wave A.

The ADAMS-Wave A contains the CDR scale (Clinical Dementia Rating) to assess and quantify severity of dementia based on its effects on various areas of functioning including: (1) memory, (2) orientation, (3) judgment and problem solving, (4) community affairs, (5) home and hobbies, and (6) personal care, into the following seven scores1:

1. 0: No dementia
2. 0.5: Mild cognitive impairment or questionable/very mild dementia
3. 1: Mild dementia
4. 2: Moderate dementia
5. 3: Severe dementia
6. 4: Profound dementia
7. 5: Terminal dementia

In this package, the CDR scale is used to create a severity measure.

# Methods

To create a severity of dementia measure for the HRS sample, the followings steps were conducted:

1. We regrouped CDR categories for regression models based on the distribution of CDR in ADAMS
2. We modeled the independent variables that explain CDR in ADAMS
3. We used the model estimates to predict a severity score for the HRS sample

Step 1) Table 1 shows the ADAMS sample characteristics by categories of CDR. Table 2 includes a complete list of variables used in the analyses, and describes how variables were recoded and how the missing observations were handled/filled. To fill out some of the missing observations in the ADAMS sample, a wave matching was done between the ADAMS and HRS sample. First, the date difference between HRS waves and ADAMS interview date was calculated, second the absolute value of these time differences was taken and the minimum absolute value was selected to match with the closest waves of HRS. As a result, the ADAMS sample was merged with the closest waves from HRS, waves 5 and 6, to fill out the missing values for some of the variables.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 1. Characteristics of ADAMS sample, wave A (N=852) | | | | | | |
| Variables | **Clinical Dementia Rating (CDR)** | | | | | |
|  | *none (0)* | *mild impairment (0.5)* | *mild (1.0)* | *moderate (2.0)* | *severe (3.0)* | *profound & terminal (4.0 & 5.0)* |
| N | 228 | 365 | 108 | 66 | 58 | 27 |
| **Age - (Q1, median, Q3)** | (73, 76, 80) | (70, 74, 83) | (78, 82, 86) | (77, 86, 90) | (81, 84, 87) | (83, 86, 88) |
| **Age -wave A** |  |  |  |  |  |  |
| 70 - 79 | 139(70) | 158(53) | 30(33) | 14(28) | 8(14) | \* |
| 80 - 89 | 80(27) | 159(39) | 54(53) | 27(38) | 32(68) | 15(79) |
| 90+ | 9(3) | 48(8) | 24(14) | 25(34) | 18(18) | 10(18) |
| **Gender** |  |  |  |  |  |  |
| male | 107(40) | 173(42) | 42(45) | 16(21) | 12(22) | \* |
| female | 121(60) | 192(58) | 66(55) | 50(79) | 46(78) | 23(93) |
| **Education years - (Q1, median, Q3)** | (11, 12, 14) | (9, 11, 12) | (6, 9, 12) | (7, 11, 14) | (9, 11, 14) | (8, 11, 12) |
| **Education years** |  |  |  |  |  |  |
| Grade 0 - 11 | 73(24) | 216(41) | 73(55) | 36(40) | 26(40) | 14(34) |
| Grade 12 - 15 | 112(54) | 119(48) | 30(42) | 21(44) | 27(47) | 11(61) |
| Grade 16+ | 43(21) | 30(11) | \* | 9(16) | \* | \* |
| **Race ethnicity** |  |  |  |  |  |  |
| White/Caucasian, Not Hispanic | 190(90) | 224(80) | 65(74) | 47(82) | 47(89) | 16(47) |
| Black or African American, Not Hispanic | 19(4) | 84(10) | 30(17) | 10(8) | 9(9) | 6(8) |
| Hispanic | 15(3) | 49(8) | 10(5) | 7(8) | \* | \* |
| Other, Not Hispanic | \* | 8(1) | \* | \* | \* | \* |
| **Proxy interview** |  |  |  |  |  |  |
| yes | 8(4) | 42(9) | 36(25) | 40(60) | 47(74) | 25(94) |
| no | 220(96) | 323(91) | 72(75) | 26(40) | 11(26) | \* |
| **TICS cog measure (0-25) - (Q1, median, Q3)** | (13, 15, 18) | (8, 11, 14) | (4, 7, 9) | (0, 1, 1) | (0, 0, 2) | (0, 0, 2) |
| missing, n=39 | \* | \* | \* | \* | 17(24) | 20(78) |
| **Cognition status** |  |  |  |  |  |  |
| Dementia | 9(1) | 101(10) | 68(41) | 51(71) | 51(89) | 23(90) |
| Cognitive impairment | 49(14) | 153(33) | 33(46) | 9(20) | \* | \* |
| Normal | 170(86) | 107(56) | \* | \* | \* | \* |
| **IADL - (Q1, median, Q3)** | (0, 0, 0) | (0, 0, 0) | (0, 1, 2) | (2, 4, 4) | (3, 4, 5) | (2, 4, 4) |
| **IADL** |  |  |  |  |  |  |
| 0 | 201 (88) | 215 (66) | 21 (32) | \* | \* | \* |
| 1 to 2 | 21 (11) | 90 (23) | 37 (31) | \* | \* | \* |
| 3 to 5 | 6 (2) | 60 (11) | 50 (37) | 63 (95) | 56 (99) | 26 (98) |
| **ADL - (Q1, median, Q3)** | (0, 0, 0) | (0, 0, 0) | (0, 0, 1) | (1, 2, 4) | (3, 4, 4) | (4, 4, 5) |
| **ADL** |  |  |  |  |  |  |
| 0 | 191 (84) | 244 (69) | 48 (57) | 11 (17) | \* | \* |
| 1 to 2 | 15 (7) | 63 (14) | 36 (27) | 22 (25) | 7 (14) | \* |
| 3 to 5 | 22 (9) | 58 (17) | 24 (16) | 33 (57) | 49 (84) | 26 (99) |
| **Depression** |  |  |  |  |  |  |
| No | 216 (94) | 319 (89) | 83 (85) | 44 (62) | 30 (49) | \* |
| Yes | 12 (6) | 44 (11) | 21 (14) | 12 (14) | \* | \* |
| missing, n=57 | \* | \* | \* | 10 (24) | 26 (44) | 22 (80) |
| Some frequencies are starred in accordance with DUA (n<5) | | | | | | |

Based on the distribution observed in table 1, we decided to combine CDR categories 2 and 3, and categories 4 and 5 to address small sample sizes and facilitate distinguishing no dementia and mild dementia from dementia at later stages that may represent a delayed or late diagnosis. Thus, the measure of severity ranges from 0 to 4 as follows: 0 (CDR=0); 1 (CDR=0.5); 2 (CDR-=1); 3 (CDR=2 or CDR=3); and 4 (CDR=4 or CDR=5).

| Table 2. List of variables | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| HRS Variable name | Label | ADAMS Variable name | Recoded? | How recoded? | Missing observations?  If yes, how handled? |
| CDR | Clinical Dementia Rating Scale | ADCDRSTG | Yes | 0 (ADCDRSTG=0)  1 (ADCDRSTG=0.5)  2 (ADCDRSTG-=1)  3 (ADCDRSTG=2 **AND** ADCDRSTG=3)  4 (ADCDRSTG=4 **AND**  ADCDRSTG=5) | No |
| Age | Subject age at time of selection | AAGE | No | N/A | No |
| Gender | Subject gender | GENDER | No | N/A | No |
| EDYRS | Subject years of formal education | EDYRS | Yes | 1 (Grade 0-11)  2 (Grade 12-15)  3 (Grade 16 and above) | No |
| Racethn | Race and ethnicity | RACE & ETHNIC | Yes | 1 (White, non-Hispanic)  2 (Black, non-Hispanic)  3 (Hispanic)  4 (Other, non-Hispanic) | No |
| Cogstate | Cognition status |  | Yes | 1 (Dementia)  2 (Cognitive impairment)  3 (Normal) | No |
| ADL |  |  | Yes | ADL is calculated by summing below variables:  AGQ30A: Getting across room  AGQ30B: Dressing  AGQ30C: Bathing  AGQ30D: Eating  AGQ30E: Getting out of bed  ADL is categorized into 3 categories:  0 (ADL=0)  1 (ADL=1,2)  2 (ADL=3,4,5) | Yes, filled by matching to the closest HRS wave. |
| IADL |  |  | Yes | IADL is calculated by summing below variables:  AGQ30G: Preparing a meal  AGQ30H: Shopping for groceries  AGQ30I: Making telephone calls  AGQ30J: Taking medication  AGQ30K: Managing money  IADL is categorized into 3 categories:  0 (IADL=0)  1 (IADL=1,2)  2 (IADL=3,4,5) | Yes, filled by matching to the closest HRS wave. |
| CIDI\_CESD | Depression | Ciditot | Yes | 0 (ciditot of 0 to 4 **OR** CESD of 0 to 2)  1 (ciditot score of 5 to 7 **OR** CESD of 3 to 8)  2 (missing) | Yes, filled by matching to the closest HRS wave. |
| Proxy | HRS interview proxy status | PROXY | No | N/A | No |

Step 2) First, we modeled the value of CDR in the ADAMS sample based on variables available to all HRS respondents in order to extrapolate the relationship onto the entire longitudinal HRS sample later. We predicted CDR using age, gender, years of education, race/ethnicity, proxy interview status, cognition status, ADL, IADL, and depression and whether a proxy responded for the subject. We experimented with multiple models, including OLS, Poisson, Negative Binomial, Zero-inflated Negative Binomial, and Logistic regression and selected the Poisson model based on overall prediction accuracy and robustness. The regression results are shown below in table 3.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 3. Poisson model predicting CDR (0,1,2,3,4) among ADAMS sample, N=852 | | | | | |
| *Parameter* | *Estimate* | *SE* | *95% CI* | | *p-value* |
| Intercept | -5.2210 | 3.6755 | -12.4249 | 1.9829 | 0.1555 |
| Age | 0.0911 | 0.0871 | -0.0796 | 0.2618 | 0.2956 |
| Age squared | -0.0005 | 0.0005 | -0.0015 | 0.0005 | 0.3441 |
| Female | 0.0375 | 0.0754 | -0.1103 | 0.1853 | 0.6192 |
| Education |  |  |  |  |  |
| HS to some college | 0.0636 | 0.0803 | -0.0938 | 0.2209 | 0.4286 |
| college graduate and above | 0.0929 | 0.1343 | -0.1702 | 0.3560 | 0.4890 |
| Race/ ethnicity |  |  |  |  |  |
| Black, non-Hispanic | 0.0253 | 0.0923 | -0.1556 | 0.2062 | 0.7840 |
| Hispanic | -0.0021 | 0.1245 | -0.2461 | 0.2419 | 0.9867 |
| Others, non-Hispanic | 0.0801 | 0.2007 | -0.3134 | 0.4735 | 0.6899 |
| Interview proxy status |  |  |  |  |  |
| Proxy | 0.1913 | 0.0938 | 0.0074 | 0.3752 | 0.0415 |
| Cognition status |  |  |  |  |  |
| has dementia | 0.8206 | 0.1342 | 0.5576 | 1.0836 | <.0001 |
| cognitive impairment | 0.5946 | 0.1232 | 0.3532 | 0.8359 | <.0001 |
| ADL |  |  |  |  |  |
| ADL score (1,2) | 0.1201 | 0.1121 | -0.0997 | 0.3399 | 0.2844 |
| ADL score (3,4,5) | 0.1981 | 0.1114 | -0.0202 | 0.4165 | 0.0753 |
| IADL |  |  |  |  |  |
| IADL score (1,2) | 0.1568 | 0.1245 | -0.0871 | 0.4008 | 0.2077 |
| IADL score (3,4,5) | 0.4563 | 0.1252 | 0.2110 | 0.7016 | 0.0003 |
| Depression |  |  |  |  |  |
| have depression score (5,6,7) | -0.0957 | 0.1127 | -0.3166 | 0.1252 | 0.3957 |
| missing | 0.2848 | 0.1104 | 0.0684 | 0.5011 | 0.0099 |

Step 3) We then used the models estimates in table 3 to predict a severity score for all HRS wave 6 respondents. We then compared the distribution of our predicted severity score with the distribution of original severity score, CDR, to assess how well we re-predicted the CDR in ADAMS sample, as well as in the HRS wave 6 sample. Table 4 shows the distribution results.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4. Comparison of distribution of observed and predicted severity measure across different models in both samples of ADAMS wave A (N=852), and HRS wave (N=7634) | | | | | | | | | | | | | |
|  | **OBSERVED** | | | | | | | | | | | | |
|  | | *mean* | *SD* | *min* | *p1* | *p5* | *p10* | *p25* | *p50* | *p75* | *p90* | *p95* | *max* |
| Categorical CDR (0,1,2,3,4) | | 1.25 | 1.1 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 3 | 4 |
| Binary CDR (0,1) | | 0.73 | 0.44 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
|  | **PREDICTED** | | | | | | | | | | | | |
|  |  | *mean* | *SD* | *min* | *p1* | *p5* | *p10* | *p25* | *p50* | *p75* | *p90* | *p95* | *max* |
| OLS | ADAMS | 1.25 | 0.91 | 0.05 | 0.14 | 0.23 | 0.29 | 0.50 | 0.97 | 1.89 | 2.54 | 3.20 | 3.59 |
| HRS-W6 | 0.83 | 0.84 | 0.00 | 0.08 | 0.16 | 0.20 | 0.29 | 0.48 | 1.02 | 2.05 | 3.18 | 3.63 |
| Poisson | ADAMS | 1.25 | 0.92 | 0.30 | 0.32 | 0.34 | 0.36 | 0.44 | 0.88 | 1.83 | 2.66 | 3.17 | 3.93 |
| HRS-W6 | 0.8 | 0.78 | 0.28 | 0.30 | 0.32 | 0.33 | 0.36 | 0.44 | 0.83 | 1.73 | 3.07 | 4.26 |
| Zero Inflated | ADAMS | 1.24 | 0.92 | 0.04 | 0.11 | 0.24 | 0.35 | 0.46 | 0.91 | 1.83 | 2.64 | 3.09 | 4.26 |
| HRS-W6 | 0.78 | 0.8 | 0.01 | 0.03 | 0.17 | 0.19 | 0.36 | 0.46 | 0.84 | 1.7 | 3.06 | 4.5 |
| Ordered probit | ADAMS | 1.19 | 1.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 3.00 | 3.00 | 4.00 |
| HRS-W6 | 0.68 | 0.92 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 2.00 | 3.00 | 4.00 |
| Multinomial | ADAMS | 1.19 | 1.06 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 2.00 | 3.00 | 3.00 | 4.00 |
| HRS-W6 | 0.72 | 0.9 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 2.00 | 3.00 | 4.00 |
| Logistic | ADAMS | 0.73 | 0.28 | 0.12 | 0.22 | 0.24 | 0.28 | 0.47 | 0.86 | 0.98 | 1.00 | 1.00 | 1.00 |
| HRS-W6 | 0.58 | 0.3 | 0.12 | 0.18 | 0.22 | 0.23 | 0.28 | 0.54 | 0.91 | 1.00 | 1.00 | 1.00 |

Figure 1 below depicts the observed CDR and predicted severity measure from the Poisson model (final chosen model) in the ADAMS sample to show how well we have predicted severity. In this figure, values for predicted severity are bucketed (rounded) into the nearest 0.25th place. We censor buckets with a frequency less than 5.

Figure 1. Observed and predicted CDR score in the ADAMS sample

Reference:

1. ADAMS Supplement to the Health and Retirement Study, public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (2007)