OXFORD
UNIVERSITY PRESS

Archives
of
CLINICAL
NEUROPSYCHOLOGY

Symbol Digit Modalities Test: Regression-Based Normative Data and Clinical Utility

Robert P. Fellows, Maureen Schmitter-Edgecombe*

Department of Psychology, Washington State University, WA, USA

*Corresponding author at: Department of Psychology, Washington State University, Pullman, WA 99164-4820, USA. Tel.: +1-509-335-0170; fax: +1-509-335-5043. *E-mail address*: schmitter-e@wsu.edu

Editorial Decision 19 March 2019; Accepted 22 March 2019

Abstract

Objective: The purpose of this study was to provide regression-based normative data for the written, oral, and incidental recall trials of the Symbol Digit Modalities Test (SDMT).

Method: Regression-based normative equations for the written and oral trials were derived from 536 healthy men and women between the ages of 18 and 91. Normative equations for the incidental recall trial are provided for a subset of the normative sample (age range = 60–91). The clinical utility of the newly developed norms was examined by comparing mean performance and rates of impaired scores for participants with traumatic brain injury (TBI), mild cognitive impairment (MCI), and dementia. Within-group analyses were used to compare the new norms to the original published norms.

Results: Age, education, and sex were all significant predictors of written trial performance, age and education were significant predictors of oral trial performance, and only age predicted incidental recall trial performance. As expected, the TBI group demonstrated the highest rates of impaired performance on both written and oral trials. Participants with dementia showed the highest rate of impaired scores on the incidental recall trial, followed by participants with amnestic MCI. Compared to traditional norming methods, the regression-based norms classified more clinical participants as impaired on both the written and oral trials.

Conclusions: Comprehensive regression-based normative equations with demonstrated clinical utility are provided to improve the detection of cerebral dysfunction using the SDMT. A calculator with the normative equations is provided so that raw scores can be easily converted to demographically-corrected standardized scores.

Keywords: Norms; Assessment; Processing speed; Memory; Symbol substitution test; Mild cognitive impairment; Traumatic brain injury; Dementia; Norms calculator

Normal aging is associated with a gradual decline in information processing speed (Salthouse, 2000). Therefore, methods to parse the effects attributable to normal aging from those associated with pathological aging are essential to accurate diagnosis and tracking of cognitive decline. As such, demographically-corrected normative data are used to estimate expected level of cognitive functioning based on relevant demographic characteristics. Traditional methods for developing norms require large sample sizes. This is because demographic variables are divided into discrete categories (e.g., age bands, binary education classifications), which limits the precision of cognitive estimates. In contrast, regression-based norms have the advantage of requiring much smaller sample sizes (Oosterhuis, van der Ark, & Sijtsma, 2016), and improving diagnostic precision (Burggraaff, Knol, & Uitdehaag, 2017).

The Symbol Digit Modalities Test (SDMT; Smith, 1982) is a digit substitution test that requires divided attention, perceptual speed, visual scanning speed, and tracking and is very sensitive to various neurological disorders and acquired brain injury (Strauss, Sherman, & Spreen, 2006). For this task, the examinee is instructed to quickly match a symbol to its corresponding number based on the way it is matched in the key at the top of the page. The SDMT differs from other substitution

doi:10.1093/arclin/acz020

tests in that it includes two modalities of administration, written and oral. The advantage of the oral modality is that it is not susceptible to the potentially confounding influence of motor functioning, which is particularly relevant to obtaining accurate assessment of neurocognitive abilities in some clinical populations. For example, an individual with essential tremor may perform in the impaired range on the written version and in the intact range on the oral version, likely due to motor dysfunction, rather than cognitive impairment. For this reason, clinicians and researchers tend to use either the oral or written version, depending on the characteristics of the examinee. Administration of both the written and oral testing versions, however, may have clinical utility. For example, performance discrepancies may provide useful information about capacity to learn information over the two trials and comparisons between motor and non-motor speeded tasks.

To assess automatic learning of information, Uchiyama and colleagues (1994) developed an incidental recall trial for the SDMT, administered immediately following the standard administration of the written trial. Prior research has shown that incidental learning on a similar digit substitution test (i.e., WAIS-R Digit Symbol subtest) moderately correlated with delayed recall on a verbal memory test and a visual memory composite that included the Rey-Osterrieth Complex Figure delayed recall and WMS-R Visual Reproduction delayed recall, but did not significantly correlate with graphomotor speed, visual scanning speed, or either trial of the Trail Making Test (Joy, Kaplan, & Fein, 2003). This finding suggests that the incidental recall trial may be useful for assessing automatic encoding of visual information.

Despite the sensitivity of the SDMT to detect cognitive change associated with numerous neurological conditions, normative data for all three SDMT trials (i.e., written, oral, incidental recall) is relatively limited. Although SDMT normative data derived from large samples (>3000) have been published, norms are only provided for either the written version (Kiely, Butterworth, Watson & Wooden, 2014; Uchiyama et al., 1994) or the oral version (Jorm, Anstey, Christensen & Rodgers, 2004), not both. Data for the original norms from the SDMT manual (Smith, 1982) appear to have been collected in an unstandardized and nonsystematic manner and therefore other normative sources have been recommended (Strauss et al., 2006). Other norms for administration of both written and oral trials are available (Sheridan et al., 2006; Yeudall, Fromm, Reddon, & Stefanyk, 1986) but the utility of these data are limited due to the relatively circumscribed age ranges (age < 50). Of the two studies that have published normative data for the incidental recall trial (Manly et al., 2011; Uchiyama et al., 1994), both samples were relatively young (i.e., mean age less than 40) which limits application to older populations who are more likely to experience memory deficits from age-related neurodegenerative processes.

All of the aforementioned studies have also used a conventional approach to normative data, providing data for discrete age bands (e.g., ages 35–39, 40–44). These data are adequate when raw scores are normally distributed and cell sizes are large enough to provide stable estimates of performance. However, in addition to age, other demographic factors, such as sex and level of education, can influence performance on neuropsychological tests. Using a conventional normative approach to account for these additional demographic variables, discrete classifications require dividing groups into increasingly smaller categories (e.g., women ages 50–54 with 12 or more years of education), reducing cell sample size, which can reduce estimate stability. As an alternative, continuous norms can be developed using multiple regression equations. In this method, multiple regression equations are used to estimate an individual's predicted level of performance based on a set of variables (e.g., age, years of education, sex). The difference between the demographically predicted score and actual score is then standardized and used for clinical interpretation.

Compared to conventional discrete norms, regression-based normative data for the oral version of the SDMT derived from relatively small groups of healthy controls (i.e., ≤100) was found to adequately account for demographic influences in several studies (Berrigan et al., 2014; Burggraaff et al., 2017; Parmenter, Testa, Schretlen, Weinstock-Guttman, & Benedict, 2010). For example, Parmenter and colleagues (2010) found that compared to discrete norms, regression-based norms classified significantly more patients with multiple sclerosis as impaired on tests of visual memory (i.e., BVMT), speeded processing (i.e., SDMT) and executive functioning (D-KEFS Sorting). Similarly, Burggraaff and colleagues (2017) found that regression-based norms more accurately detect abnormal cognition in multiple sclerosis, compared to conventional norming methods. Findings from these studies underscore the utility of regression-based norms to more precisely determine an individual's estimated level of performance on cognitive tests.

The first goal of the current investigation is to develop demographically-corrected normative data for the SDMT using a multivariate regression-based approach. The second study goal is to examine the clinical utility of the newly developed normative data in TBI, MCI, and dementia. We expect to observe incremental rates of impaired test performance between healthy adults, MCI, and dementia. In addition, we expect that TBI participants will show the highest rates of impairment on the SDMT. Our third goal is to compare previously published norms from the SDMT manual (Smith, 1982) to the newly developed norms. We hypothesize that the regression-based norms will be more sensitive to cognitive impairment.

Method

Participants

Participants were recruited through advertisements, word-of-mouth, health fairs, community presentations, and physician referrals. Younger adults were also recruited through the university research subject pool. Participants underwent a standardized clinical interview and neuropsychological evaluation as a part of several different research studies completed in a university research laboratory. Only healthy controls from these studies were included in the normative sample. Demographic information and medical history were collected through a standardized clinical interview. Exclusion criteria for the normative sample included history of neurological disorder, learning disability, traumatic brain injury or stroke with residual deficit, uncorrected visual impairment, mild cognitive impairment, dementia, or presence of psychiatric disorder with known cognitive effects (e.g., bipolar disorder, schizophrenia, severe depression). After applying exclusion criteria, 536 participants ranging from 18 years to 91 years of age were included in the normative sample. The sample included more women (70.3%) than men (29.7%), was mostly non-Hispanic white (88.6%), and years of education completed ranged from 9 to 20 (M = 15.7, SD = 2.6). All participants provided informed consent and these studies were approved by the local IRB.

Participants with MCI (n = 128) and dementia (n = 34) were included to examine the utility of the newly developed norms. Participants were classified with MCI using established criteria (Petersen & Morris, 2005; Petersen et al., 2001). These criteria included self- or informant report of subjective memory impairment for at least 6 months, scoring at least 1.5 standard deviations below age-matched norms or relative to prior testing on a measure in one or more cognitive domains (i.e., memory, language, executive functioning, and/or speeded processing), generally preserved functional abilities, and not meeting Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition (DSM–IV) criteria for dementia (American Psychiatric Association, 2000). Of note, the majority of participants (79.7%) who met criteria for MCI had the amnestic subtype (n = 102). All of the non-amnestic MCI participants (n = 26) were impaired on at least one test other than the SDMT and also met MCI criteria. Therefore, MCI classification was not due to performance on the SDMT. Participants were classified with dementia using DSM-IV diagnostic criteria.

Participants with traumatic brain injury (TBI) were also included to test the clinical utility of the newly developed norms. TBI participants (n = 74) were recruited from consecutive admissions to an inpatient rehabilitation program. The Glasgow Coma Scale (GCS; Teasdale & Jennett, 1974) was used to assess TBI severity and was administered at the scene of the accident or in the emergency department. Of the 74 TBI participants, two were classified as mild (GCS = 13 or high), 28 were classified as moderate based on a GCS between 9 and 12 (n = 12) or a GCS of higher than 12 with positive neuroimaging findings (n = 16), and 43 were classified as severe (GCS = 8 or less). One participant was excluded for a missing GCS score. These participants were tested after emergence from posttraumatic amnesia as determined by the Galveston Orientation and Amnesia Test or clinical interview. Participant characteristics are presented in Table 1.

Estimates of intellectual ability for the entire sample were derived from various measures (i.e., Weschler Test of Adult Reading, Shipley Institute of Living Scale – Vocabulary subtest, and Barona method), depending on data available. Percentile cut-offs for standard scores or T-scores were used to classify level of premorbid functioning into three categories: Low Average or below (≤ 24th percentile), Average (25th – 74th percentile), High Average or above (≥75th percentile). As shown in Table 1, premorbid intellectual ability for the majority of individuals in the normative sample, as well as the MCI and dementia groups, was estimated to be at or above average. Whereas, the majority of TBI participants fell in the average to below average range of estimated premorbid intellectual functioning.

Materials

Symbol Digit Modalities Test *(Smith, 1982).* Both written and oral versions of the SDMT were administered according to the standardized instructions. At the top of the SDMT form is a key with boxes containing nine unique geometric shapes paired with numbers ranging from 1 to 9. Below the key are rows of boxes with geometric shapes in the top boxes with rows of empty boxes. The participant is instructed to fill each empty box with the number that matches the shape using the key at the top of the page. To start, the participant completes 10 practice items. Following the practice trial, the participant is then instructed to fill as many boxes as possible in 90 s, matching numbers to the corresponding shape. Immediately after completion of the written trial, the oral trial is administered. In the oral trial, rather than writing the numbers that match the shapes, the participant says them to the examiner who writes them down. Consistent with standard procedures (Smith, 1982), there was no practice trial for the oral version because it was administered after the written trial. The score for each trial is the total number of correctly paired items (maximum = 110).

Table 1. Demographic characteristics and SDMT raw scores for normative and clinical groups

Variable	Normative (n = 536) M (SD) or %	TBI (n = 73) M (SD) or %	MCI (n = 128) M (SD) or %	Dementia (n = 34) M (SD) or %
Age	52.5 (22.1)	38.1 (16.3)	71.3 (9.14)	74.0 (8.9)
Education	15.7 (2.6)	13.6 (2.4)	15.7 (2.9)	15.8 (3.1)
Sex (% female, n)	70.3 (377)	26.0 (19)	57.8 (74)	35.3 (12)
Race/ethnicity (%, n)				
Non-Hispanic White	88.6 (475)	93.2 (68)	96.1 (123)	100 (34)
Asian	3.0 (16)	—	—	—
Hispanic/Latinx	2.8 (15)	4.1 (3)	0.8 (1)	—
African American	1.5 (8)	—	0.8 (1)	—
American Indian	0.4 (2)	2.7 (2)	—	—
Pacific Islander	0.4 (2)	—	—	—
Multiracial	0.7 (4)	—	—	—
Other/no response	2.6 (14)	—	2.3 (3)	—
Premorbid IQ Estimate (%, n)				
≤ Low Average	5.3 (28)	37.7 (23)	17.5 (22)	12.9 (4)
Average	35.2 (186)	45.9 (28)	38.1 (48)	54.8 (17)
≥ High Average	59.5 (315)	16.4 (10)	44.4 (56)	32.2 (10)
SDMT Written raw	52.62 (10.89)	36.9 (12.8)	38.42 (10.44)	24.94 (10.47)
SDMT Oral raw	61.53 (14.08)	44.4 (14.1)	43.60 (11.64)	27.53 (11.91)
SDMT Recall raw	11.61 (3.70)	8.8 (4.7)	6.64 (4.44)	1.41 (1.86)

Note: TBI = traumatic brain injury; MCI = mild cognitive impairment; SDMT = Symbol Digit Modalities Test. IQ = Intelligence Quotient.

Incidental recall. Directly following the SDMT oral trial, the participant is presented with the incidental recall trial (Uchiyama et al., 1994). The incidental recall trial consists of a line of 15 boxes (the last line of the SDMT standard form) in which each of the nine symbols are included at least once. The participant is instructed to fill in the numbers that match each of the symbols. The number of correctly paired items is the total score (maximum = 15). This scoring differs from the method originally described by Uchiyama and colleagues (1994). In the original scoring method, for cases in which the same symbol appears twice, the participant is given credit if they correctly pair the symbol in one instance even if they incorrectly pair it in another. For the purposes of the current study, each item was scored individually to minimize capitalizing on chance.

Statistical analyses. Multiple regression analyses (MRA) were conducted to develop regression-based norms (RBNs) using previously described procedures (Parmenter et al., 2010; Shirk et al., 2011; Testa, Winicki, Pearlson, Gordon & Schretlen, 2009). First, raw scores from the 536 participants who met inclusion criteria for the normative sample were normalized using Blom's formula (Blom, 1958). This method uses the cumulative proportion estimates of the normal distribution to rank cases using the following formula:

$$z = \frac{r - 3/8}{w + 1/4}$$

z = z-score
w = sum of the case weights
r = rank

Then resulting z-scores were converted to scaled scores (M = 10, SD = 3; see Table 2).

MRA were used to estimate predicted performance based on demographic characteristics (age, sex, and education). Results of the regression analyses were used to determine the amount of variance explained by each independent variable, and then these equations were applied to predict individual performance. SDMT scaled scores were regressed on demographic variables (age, sex, and education). To reduce potential multicollinearity among predictor variables (Burggraaff et al., 2017), the age variable was centered by subtracting the sample mean (M_{age} = 52.509328) and the centered variable was then squared. To account for possible nonlinear effects on performance, both age-centered (agec) and age-centered squared (agec^2) variables were entered into the regression equations. Interaction terms were created for all predictors and were tested in the regression equations. None of the interaction terms were significant in any of the models, so they were not retained. Equations derived from the normative sample were used to calculate demographically-corrected z-scores (actual scaled score – predicted scaled score/RMSE). Z-scores for the SDMT oral, written, and recall trials were converted into T-scores (M = 50, SD = 10).

Table 2. Raw score- to scaled-score conversions

	SDMT Raw Scores		
Scaled Score	Written	Oral	Incidental Recall (ages 60–91)
1	≤ 25	≤ 28	
2	26–27	29–31	0–1
3	28–30	32–34	
4	31–34	35–40	2–3
5	35–37	41–42	4
6	38–40	43–46	5–6
7	41–44	47–49	7
8	45–47	50–53	8
9	48–50	54–57	9–10
10	51–53	58–62	11
11	54–57	63–66	12
12	58–60	67–74	13–14
13	61–66	75–78	15
14	67–69	79–83	
15	70–74	84–90	
16	75–80	91–99	
17	81–83	100–103	
18	≥84	≥104	

Regression analyses included agec, agec^2, education (years), and sex (1 = female, 2 = male). Predicted scores were calculated using the following equations based on the regression equations from the normative group:

$$\text{Scaledscore}_{\text{predicted}} = \text{constant} + B_{\text{agec}}(\text{agec}) + B_{\text{agec}^2}(\text{agec}^2) + B_{\text{edu}}(\text{edu}) + B_{\text{sex}}(\text{sex})$$

Normative equations were applied to participants in each of the clinical groups (i.e., MCI, dementia, TBI) by first subtracting the predicted scaled score, using equations derived from the normative sample, from the actual scaled score and then dividing the difference score by the root mean squared error (RMSE). To obtain an overall index of performance, the mean T-scores for both the written and oral trials were used to produce a composite score. Specifically, a z-score was calculated from the mean of the demographically-corrected written and oral T-scores (i.e., mean (SDMT Written T-score, SDMT Oral T-score)–sample mean/sample SD) and the resulting z-score was converted into a standard score (M = 100, SD = 15). Cronbach's alpha was used to determine internal consistency of the composite score. After applying normative equations, mean level of performance and rates of impaired scores (1 SD and 1.5 SD) were calculated for each group. Additionally, to examine the utility of the incidental recall trial paired-sample t-tests were conducted to examine SDMT performance differences between MCI subtypes (i.e., amnestic and non-amnestic).

Finally, the newly developed regression-based norms were compared to the original discrete manual norms published by Smith (1982). Means and standard deviations provided in the manual grouped by age bands (i.e., 18–24, 25–34, 35–44, 45–54, 55–64, 65–78) and separated by education (i.e., ≤12 years or ≥13 years) were used to calculate z-scores which were then transformed into T-scores. Paired-sample t-tests were used to compare the performance of normative data sources (i.e., SDMT manual norms vs. regression-based norms) within each group (i.e., healthy, TBI, MCI, dementia). Correlational analyses or t-tests, when appropriate, were then used to examine residual demographic associations with scores from each of the normative sources (i.e., manual vs. regression-based). Given that the SDMT manual only provides normative data for individuals up to age 78, only a subset of participants (i.e., ages 18–78; n = 679) were included in the analyses comparing normative methods.

Results

Results of the regression analyses are presented in Table 3. Outliers did not significantly influence any of the measures (Cook's Ds < 1). Visual inspection of the plotted residuals revealed normality for written and oral scores, with no evidence of heteroscedasticity. Scores for incidental recall trial were not normally distributed. Further examination of the data revealed considerable ceiling effects, particularly for younger age groups (i.e., ages 18–24, 77%; 25–59, 46%; 60–91, 18%). Therefore, regression-based norms for the incidental recall trial were only developed for older adults (ages 60–91; n = 261).

Table 3. Regression coefficients for SDMT scores in the normative sample ($n = 536$)

Variables	B	SE B	β	t
SDMT Written				
(Constant)	9.194523	0.714837		12.86***
Agec	−0.105976	0.005125	−0.774274	−20.68***
Agec^2	−0.001265	0.000270	−0.173926	−4.69***
Education	0.125023	0.040867	0.106522	3.06**
Sex	−0.471566	0.215039	−0.071411	−2.19*
RMSE	2.22686			
R^2	0.455984***			
SDMT Oral				
(Constant)	9.110394	0.721977		12.62***
Agec	−0.101047	0.005177	−0.745949	−19.52***
Agec^2	−0.001041	0.000273	−0.144540	−3.82***
Education	0.114401	0.041275	0.098487	2.77**
Sex	−0.292926	0.217187	−0.044821	−1.35
RMSE	2.249101			
R^2	0.437677***			
SDMT Recall (only ages 60–91; $n = 261$)				
(Constant)	9.207524	1.022452		9.01***
Agec	−0.090098	0.025023	−0.237369	−3.60***
Agec^2	0.001275	0.002807	0.030009	0.45
Education	−0.012126	0.058539	−0.012771	−0.21
Sex	0.271714	0.404478	0.041575	0.67
RMSE	2.653927			
R^2	0.052475**			

Note: SDMT = Symbol Digit Modalities Test; RMSE = Root mean square error. ***$p < .001$; **$p < .01$; *$p < .05$.

The age variable was re-centered ($M = 71.245211$) for the regression equation predicting incidental recall (raw score: $M = 10.0$, SD = 3.7) in the subset of older adults.

Demographic Effects

Regression models for each of the SDMT variables are presented in Table 3. The model predicting SDMT written performance accounted for 45.6% of the variance, $R^2 = .456$, $F(4, 531) = 113.11$, $p < .001$. Age, sex, and education all were significantly associated with performance on the written trial, with older age, male sex and fewer years of education predicting lower written scores. The model predicting SDMT oral performance accounted for 43.8% of the variance, $R^2 = .438$, $F(4, 531) = 103.32$, $p < .001$. Older age and fewer years of education significantly predicted lower SDMT oral trial performance, but no significant sex effect was observed. As previously noted, regression equations for the SDMT incidental recall trial were only computed for adults ages 60–91. The results of the model predicting SDMT incidental recall performance accounted for 5.2% of the variance, $R^2 = .052$, $F(4, 256) = 3.54$, $p = .008$. Only age significantly predicted incidental recall performance. Equations for calculating T-scores and standard scores are presented in the Appendix. As illustrated in Fig. 1, on average, SDMT raw scores declined with increasing age and participants performed better on the oral trial, compared to the written trial, across all age groups.

Utility of Regression-based Normative Equations

Previously described procedures for applying regression-based norms were used to convert raw scores to scaled scores to demographically-corrected T-scores for MCI, dementia, and TBI groups. The SDMT composite standard score was calculated from written and oral T-scores. The SDMT composite score showed good internal consistency in the normative sample (Cronbach's alpha = .85). Table 4 presents mean T-scores and rates of impaired scores for normative, TBI, MCI, and dementia groups. Overall, the TBI group had the lowest scores and highest rates of impairment on both the written and oral trials; whereas, the dementia group demonstrated the highest rate of impairment on the incidental recall trial. Rates of impaired incidental recall scores were not calculated for the TBI group because only 11 participants fell within the normative age range.

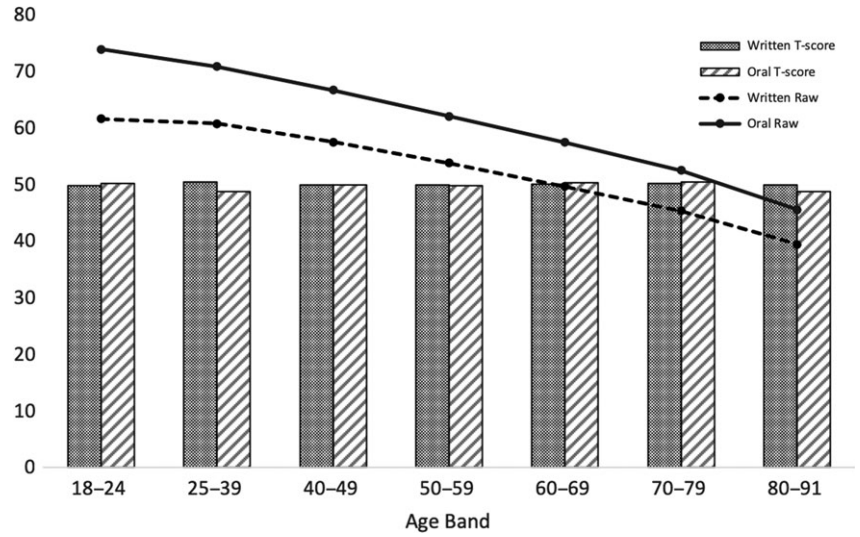

Fig. 1. Raw scores (presented as lines) and demographically-corrected T-scores (presented as bars) for the normative sample by age group.

Table 4. SDMT T-scores, standard scores, and impaired scores for normative and clinical groups

Variable	Normative (n = 536) *M* (*SD*) or %	TBI (n = 73) *M (SD)* or %	MCI (n = 128) *M* (*SD*) or %	Dementia (n = 34) *M* (*SD*) or %
SDMT Written	49.99 (9.96)	22.76 (13.45)	40.26 (11.64)	28.09 (12.07)
$T \leq 40$	15.5	90.4	50.8	85.3
$T \leq 35$	8.4	82.2	37.5	70.6
SDMT Oral	50.00 (9.96)	24.79 (12.51)	39.18 (11.55)	26.46 (10.31)
$T \leq 40$	17.2	86.3	54.7	85.3
$T \leq 35$	6.9	82.2	39.1	79.4
SDMT IR[a]	50.01 (9.94)	—	39.27 (11.50)	26.72 (6.02)
$T \leq 40$	16.5	—	56.6	100
$T \leq 35$	7.3	—	38.1	87.5
SDMT composite	100.00 (15.00)	57.69 (20.17)	83.42 (17.96)	63.35 (17.76)
SS ≤ 85	15.9	90.4	53.9	88.2
SS ≤ 75	5.0	84.9	38.3	73.5

Note: SDMT = Symbol Digit Modalities Test; TBI = traumatic brain injury; MCI = mild cognitive impairment; IR = incidental recall trial; SS = standard score. [a]For incidental recall trial, normative n = 261, MCI n = 113, dementia n = 32. Incidental recall T-scores were not computed for the TBI group because only 11 individuals fell within the normative age range.

Comparison of MCI subgroups revealed that the aMCI group scored significantly lower on the incidental recall trial ($M_{T\text{-}score}$ = 37.4, SD = 10.9) compared to the non-amnestic MCI group ($M_{T\text{-}score}$ = 46.4, SD = 10.9, t = 3.52, df = 111, p = .001), but not on any of the other SDMT indices. A higher proportion of aMCI participants (44%; 40/90) demonstrated impaired (i.e., T-score ≤ 35) SDMT incidental recall performance compared to naMCI participants (13%; 3/23), $\chi^2(1, N = 113) = 7.66$, p = .006. The proportion of impaired scores between MCI subgroups did not significantly differ on any of the other SDMT scores.

Comparison of normative standards

Paired sample t-tests were used to compare the original manual norms to the regression-based norms for the SDMT written and oral trials. Given that the original published manual norms (Smith, 1982) provide data for ages from 18 to 78 years, the same age range was used for comparisons across norming methods, which resulted in a smaller subset of participants (healthy = 489; TBI = 72; MCI n = 98; dementia n = 21). Compared to the manual norms, the regression-based T-scores were significantly lower for all three clinical groups on both SDMT trials (see Table 5).

Next, participants were classified as either intact (T-score > 35) or impaired (T-score ≤ 35) using each norming method (i.e., original manual and regression-based). The new regression-based norms classified 27% more participants with MCI as

Table 5. Results of paired samples *t*-tests comparing SDMT manual (old) to regression-based (new) T-scores

	Normative (n = 489)			TBI (n = 71)			MCI (n = 98)			Dementia (n = 21)		
Variable	Old	New	*t*	Old	New	*t*	Old	New	*t*	Old	New	*t*
Written												
Mean	53.74	50.06	−19.08*	30.64	22.70	−14.92*	48.39	41.18	−13.89*	33.51	24.99	−5.16*
(SD)	(9.04)	(9.88)		(12.34)	(13.62)		(9.90)	(12.09)		(13.34)	(13.17)	
Oral												
Mean	54.54	50.13	−23.32*	33.29	24.70	−15.96*	45.95	39.84	−12.39*	30.45	23.73	−3.19*
(SD)	(10.07)	(9.90)		(10.55)	(12.67)		(10.07)	(12.13)		(14.47)	(10.79)	

Note: SDMT = Symbol Digit Modalities Test; TBI = traumatic brain injury; MCI = mild cognitive impairment. *p < .01.

impaired on the written trial (8% vs. 37%), $\chi^2(1, N = 98) = 20.32$, $p < .001$, (95% CI 15.26 –37.11) and 24% more on the oral trial (14% vs. 38%), $\chi^2(1, N = 98) = 13.98$, $p < .001$, (95% CI 11.29–34.85). In addition, the new norms classified 17% more TBI participants as impaired on the written trial (65% vs. 82%), $\chi^2(1, N = 71) = 5.14$, $p = .02$, (95% CI 2.31–30.60), and 25% more on the oral trial (56% vs. 82%), $\chi^2(1, N = 71) = 10.63$, $p = .001$, (95% CI 10.23–39.07). Finally, the new norms classified 14% more dementia participants as impaired on the written trial (62% vs. 76%), and 19% more on the oral trial (67% vs. 86%), but the differences were not statistically significant between norming methods p's > .1.

Within the sample of healthy participants, small but statistically significant demographic effects were evident after applying the original manual norms. Specifically, written trial T-scores were associated with age, $r(487) = .161$, $p < .001$, and sex, $t(487) = 2.43$, $p = .015$, with women outperforming men. No significant effects were apparent in the oral trial. Lastly, none of the demographic variables were significantly associated with the regression-based T-scores.

Discussion

Lack of adequate normative data is a frequently reported concern of neuropsychologists (Rabin, Paolillo, Barr, 2016). Insufficient access to resources needed to collect standardization data from large samples is one of the many challenges in developing normative data using a conventional approach (e.g., stratifying by relevant demographics). In contrast, the regression-based normative approach offers the potential to provide a more accurate reference by simultaneously accounting for multiple demographic variables and requiring data from fewer individuals. Results from the current study provide regression-based demographically-corrected normative equations for the written and oral trials of the SDMT from a community-dwelling sample of cognitively intact men and women ages 18–91. Normative equations are also provided for the SDMT incidental recall trial for ages 60–91.

Demographic variables accounted for over 40% of the variance in both SDMT written and oral scores. Older age was associated with lower uncorrected scaled scores on all measures. The association between older age and lower SDMT performance is consistent with previous normative studies (Jorm et al., 2004; Kiely et al., 2014; Manly et al., 2011; Smith, 1982) and research demonstrating a decline in processing speed in later life (Salthouse, 1994, 2000). Regression-based equations may improve the precision of normative estimates by treating age as a continuous variable, rather than using discrete age band classifications. This approach also has the advantage of being able to account for nonlinear age effects (e.g., age-squared) which are often observed at the tail end of age distributions. As shown in Fig. 1, the regression-based demographic corrections in the current study provided relatively consistent estimates of performance across all age ranges. More years of education predicted higher scores on the written and oral trials, but not on the incidental recall trials. The positive association between education and SDMT oral and written trial performance is also consistent with previous studies (Jorm et al., 2004; Kiely et al., 2014; Manly et al., 2011; Smith, 1982). Similarly, prior research has shown that education is not a significant predictor of incidental recall performance (Manly et al., 2011; Uchiyama et al., 1994). In the current study, sex was only a significant predictor of SDMT written performance, with women demonstrating better performance. Relations between sex and SDMT performance has been inconsistent in prior studies. Some studies have found a significant association with sex (Jorm et al., 2004; Kiely et al., 2014), while others have not (Sheridan et al., 2006). However, when a significant effect of sex has been reported, women outperformed men (Jorm et al., 2004; Kiely et al., 2014), which is consistent with the current study results.

Another goal of this study was to evaluate the clinical utility of the newly developed norms. Within clinical groups, the rates of impaired scores remained relatively consistent across written and oral trials, with slightly more participants being classified as impaired on the SDMT composite score. The incremental rates of impaired SDMT performance across groups (e.g., dementia > MCI > healthy) provides support for the clinical utility of the newly developed norms. As expected, participants

with dementia showed the highest percentage of impaired SDMT incidental recall scores (i.e., 88%). Moreover, compared to naMCI participants, aMCI participants showed a higher rate of impaired scores on the incidental recall trial but not on either of the other SDMT trials or composite score, suggesting the incidental recall trial may assist with discriminating MCI subtypes. Consistent with prior research examining processing speed performance in TBI (Donders & Strong, 2015), high rates of impaired SDMT performance was evident in the TBI group, with over 80% of participants scoring in the impaired range on written, oral, and composite measures. The rate of SDMT composite score impairment in the TBI group was comparable to that of impaired WAIS-IV Processing Speed Index scores previously reported (74% accurately classified as impaired) in a TBI sample (Donders & Strong, 2015), providing support for the clinical utility of the SDMT composite score. Together, these findings highlight the utility of the SDMT and these regression-based norms to characterize brain dysfunction associated with both speeded processing deficits and incidental learning abilities.

When the newly developed regression-based norms were compared to previously published normative data from the SDMT manual (Smith, 1982), notable differences were observed. Within-group comparisons revealed that T-scores derived from the new normative equations differed from SDMT manual scores by over half of a standard deviation for all three clinical groups. This finding may be due to increased precision of regression-based norms to estimate expected level of performance compared to stratified norms which may be less accurate if age bands are too broad. Notably, the residual demographic associations with SDMT scores derived from the manual suggests suboptimal adjustment for these characteristics, resulting in less accurate estimates of expected performance. The discrepancies in performance between norming methods may also be due to differences in the normative samples. In the current study we excluded individuals with known MCI from the normative sample, which may have improved the sensitivity of the new norms to detect more subtle impairments.

Despite the potential advantages of using the newly developed normative equations, several limitations must be considered. The participants in the current study were mostly non-Hispanic white which limits the application of these norms to individuals with more diverse backgrounds, as race/ethnicity has been shown to influence SDMT performance (Manly et al., 2011). Although the racial/ethnic composition of the current sample is generally representative of individuals living in the region in which data for the current study was collected, additional research is needed to develop normative standards that account for the influence of race/ethnicity on test performance. Participants in the current study were also, on average, highly educated which should be considered when using these norms; however, using regression-based corrections for education may provide a more accurate estimate of expected performance than stratified norms, even in a sample with a relatively high average level of education. Moreover, the exclusion of individuals with MCI from the normative sample warrants consideration with regard to comparing relative performance (e.g., strengths and weaknesses) across a battery of tests with different normative sources. Another notable limitation is that the newly developed norms should not be used if only the oral trial is administered because exposure to the written trial likely contributes to better performance on oral trial. Similarly, the incidental recall equations should only be used when both written and oral trials are administered as learning across trials likely affects recall performance. However, the normative equations can be used if only the written trial is administered.

In summary, we provided regression-based normative data for the SDMT corrected for age, education, and sex. The norms for the written, oral, and composite scores can be used for men and women age 18–91 and incidental recall norms can be used for adults age 60–91. As the multiple steps involved in calculating the regression-based norms may increase the risk of error and could be cumbersome for clinicians to calculate, a normative calculator is provided (see Supplement). By entering examinee demographics (i.e., age, education, sex) and SDMT raw scores, the calculator automatically converts the raw scores to demographically-corrected standardized scores using embedded formulas and provides corresponding percentiles and descriptors (e.g., impaired, borderline, average, etc.). Providing support for the clinical utility of the newly developed norms, compared to the normative sample, all clinical groups demonstrated a higher rate of impairment on all of the SDMT measures, with incremental rates of impairment across MCI and dementia. In addition, comparison across clinical groups revealed that the TBI participants showed the highest rate of impairment on written and oral trials, while dementia and aMCI participants demonstrated the highest proportions of impaired performance on the incidental recall trial.

Supplementary Material

Supplementary material is available at *Archives of Clinical Neuropsychology* online.

Conflict of Interest

None declared.

Acknowledgements

NIBIB #R01 EB009675; Life Sciences Discovery Fund; NINDS #R01 NS047690.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (Rev. 4th ed.). Washington, DC: Author.

Berrigan, L. I., Fisk, J. D., Walker, L. A., Wojtowicz, M., Rees, L. M., Freedman, M. S., et al. (2014). Reliability of regression-based normative data for the oral symbol digit modalities test: An evaluation of demographic influences, construct validity, and impairment classification rates in multiple sclerosis samples. *The Clinical Neuropsychologist*, *28*, 281–299.

Blom, G. (1958). *Statistical estimates and transformed beta variables*. New York: John Wiley and Sons.

Burggraaff, J., Knol, D. L., & Uitdehaag, B. M. (2017). Regression-based norms for the symbol digit modalities test in the dutch population: Improving detection of cognitive impairment in multiple sclerosis. *European Neurology*, *77*, 246–252.

Donders, J., & Strong, C. A. H. (2015). Clinical utility of the Wechsler Adult Intelligence Scale–Fourth Edition after traumatic brain injury. *Assessment*, *22*, 17–22.

Jorm, A. F., Anstey, K. J., Christensen, H., & Rodgers, B. (2004). Gender differences in cognitive abilities: The mediating role of health state and health habits. *Intelligence*, *32*, 7–23.

Joy, S., Kaplan, E., & Fein, D. (2003). Digit symbol–incidental learning in the WAIS-III: Construct validity and clinical significance. *The Clinical Neuropsychologist*, *17*, 182–194.

Kiely, K. M., Butterworth, P., Watson, N., & Wooden, M. (2014). The Symbol Digit Modalities Test: Normative data from a large nationally representative sample of Australians. *Archives of Clinical Neuropsychology*, *29*, 767–775.

Manly, J. J., Smith, C., Crystal, H. A., Richardson, J., Golub, E. T., Greenblatt, R., et al. (2011). Relationship of ethnicity, age, education, and reading level to speed and executive function among HIV+ and HIV–women: The Women's Interagency HIV Study (WIHS) Neurocognitive Substudy. *Journal of Clinical and Experimental Neuropsychology*, *33*, 853–863.

Oosterhuis, H. E., van der Ark, L. A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, *23*, 191–202.

Parmenter, B. A., Testa, S. M., Schretlen, D. J., Weinstock-Guttman, B., & Benedict, R. H. (2010). The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society*, *16*, 6–16.

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*, 1985–1992.

Petersen, R. C., & Morris, J. C. (2005). Mild cognitive impairment as a clinical entity and treatment target. *Archives of Neurology*, *62*, 1160–1163.

Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology*, *31*, 206–230.

Salthouse, T. A. (1994). The nature of the influence of speed on adult age differences in cognition. *Developmental Psychology*, *30*, 240.

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*, 35–54.

Sheridan, L. K., Fitzgerald, H. E., Adams, K. M., Nigg, J. T., Martel, M. M., Puttler, L. I., et al. (2006). Normative Symbol Digit Modalities Test performance in a community-based sample. *Archives of Clinical Neuropsychology*, *21*, 23–28.

Smith, A. (1982). *Symbol digit modalities test* (22). Los Angeles, CA: Western Psychological Services.

Shirk, S. D., Mitchell, M. B., Shaughnessy, L. W., Sherman, J. C., Locascio, J. J., Weintraub, S., & Atri, A. (2011). A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. Alzheimer's Research & *Therapy, 3*(6), 32.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *The Lancet*, *304*, 81–84.

Testa, S. M., Winicki, J. M., Pearlson, G. D., Gordon, B., & Schretlen, D. J. (2009). Accounting for estimated IQ in neuropsychological test performance with regression-based techniques. *Journal of the International Neuropsychological Society*, *15*, 1012–1022.

Uchiyama, C. L., D'elia, L. F., Dellinger, A. M., Seines, O. A., Becker, J. T., Wesch, J. E., et al. (1994). Longitudinal comparison of alternate versions of the Symbol Digit Modalities Test: Issues of form comparability and moderating demographic variables. *The Clinical Neuropsychologist*, *8*, 209–218.

Yeudall, L. T., Fromm, D., Reddon, J. R., & Stefanyk, W. O. (1986). Normative data stratified by age and sex for 12 neuropsychological tests. *Journal of Clinical Psychology*, *42*, 918–946.

Appendix

- Use Table 2 to convert raw scores to uncorrected scaled scores.
- Center age variable (i.e., age – 52.509328) (age – 71.245211, for incidental recall)
- Square centered age variable.
- Enter variables into equations below (Note: 1 = female, 2 = male; age in years; education in years).

Normative equations for SDMT written, oral, and composite scores

SDMT Written T-score:

$$50 + 10 * [(\text{SDMT written scaled score}) - ((9.194523) - (0.105976 * \text{agec}) - (0.001265 * \text{agec}^2) + (0.125023 * \text{edu}) - (0.471566 * \text{sex}))]/2.22686$$

SDMT Oral T-score:

$$50 + 10 * [(\text{SDMT oral scaled score}) - ((9.110394) - (0.101047 * \text{agec}) - (0.001041 * \text{agec}^2) + (0.114401 * \text{edu}) - (0.292926 * \text{sex}))]/2.249101$$

SDMT Composite Standard Score:

$$100 + 15 * [((\text{SDMT written Tscore} + \text{SDMT oral Tscore})/2) - (49.99979)]/9.299655$$

Normative equation for SDMT incidental recall (ages 60–91 only)

SDMT Recall:

$$50 + 10 * [(\text{SDMT recall scaled score}) - ((9.207524) - (0.090098 * \text{agec}) + (0.001275 * \text{agec}^2) - (0.012126 * \text{edu}) + (0.271714 * \text{sex}))]/2.653927$$