Normative data for the oldest old: Trail Making Test A, Symbol Digit Modalities Test, Victoria Stroop Test and Parallel Serial Mental Operations

Katarina Fällman, Lina Lundgren, Ewa Wressle, Jan Marcusson & Elisabet Classon

To cite this article: Katarina Fällman, Lina Lundgren, Ewa Wressle, Jan Marcusson & Elisabet Classon (2019): Normative data for the oldest old: Trail Making Test A, Symbol Digit Modalities Test, Victoria Stroop Test and Parallel Serial Mental Operations, Aging, Neuropsychology, and Cognition, DOI: 10.1080/13825585.2019.1648747

To link to this article: https://doi.org/10.1080/13825585.2019.1648747
Normative data for the oldest old: Trail Making Test A, Symbol Digit Modalities Test, Victoria Stroop Test and Parallel Serial Mental Operations

Katarina Fällman, Lina Lundgren, Ewa Wressle, Jan Marcusson and Elisabet Classon

Department of Acute Internal Medicine and Geriatrics, and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden; Department of Geriatric Medicine, and Department of Social and Welfare Studies, Linköping University, Norrköping, Sweden; Department of Acute Internal Medicine and Geriatrics, and Department of Social and Welfare Studies, Linköping University, Linköping, Sweden

ABSTRACT
Normative data for evaluating cognitive function in the oldest old, aged 85 years and above, are currently sparse. The normative values used in clinical practice are often derived from younger old persons, from small sample sizes or from broad age spans (e.g. >75 years) resulting in a risk of misjudgment in assessments of cognitive decline. This longitudinal study presents normative values for the Trail Making Test A (TMT-A), the Symbol Digit Modalities Test (SDMT), the Victoria Stroop Test (VST) and the Parallel Serial Mental Operations (PaSMO) from cognitively intact Swedes aged 85 years and above. 207 participants, born in 1922, were tested at 85, 90 (n = 68) and 93 (n = 35) years of age with a cognitive screening test battery. The participants were originally recruited for participation in the Elderly in Linköping Screening Assessment. Normative values are presented as mean values and standard deviations, with and without adjustment for education. There were no clinically important differences between genders, but education had a significant effect on test results for the 85-year-olds. Age effects emerged in analyses of those participants who completed the entire study and were evident for TMT-A, SDMT, VST1 and PaSMO. When comparisons can be made, our results are in accordance with previous data for TMT-A, SDMT and VST, and we present new normative values for PaSMO.

ARTICLE HISTORY
Received 29 April 2019
Accepted 19 July 2019

KEYWORDS
Neuropsychological tests; aged; 80 and over; executive function; attention; normative

Introduction
The process of normal aging is associated with changes in cognitive function, but high age is also the main risk factor for cognitive decline due to disease (Swedish Council on Technology Assessment in Health Care, 2008). Neuropsychological tests are used to investigate whether a decline in a cognitive function is due to normal aging or a pathological process such as dementia, and the investigation relies on access to reliable normative data. Even though many of the tests used in clinical practice are...
known to be affected by normal aging, and the very old population is the fastest growing age group in Sweden and in the world (Kinsella & Wan, 2008), normative data for the oldest old (85 years of age and above) are sparse. Those data that are available are often from studies with few very old persons or from populations with large age spans, including younger elderly. To be able to give the correct diagnosis and avoid misjudgments, valid normative data for all ages are needed. This study focuses on four widely used tests where the norms for the oldest population are of varying quality and seldom include nonagenarians.

The neuropsychological tests in this study evaluate attention and executive function. Attention is the process whereby a person becomes receptive to stimuli and begins to process the information, while executive functions are the capacities that together allow for independent and purposive actions (Harada, Natelson Love, & Triebel, 2013). Executive functions are crucial for us to manage our daily life, to be able to plan ahead, to work efficiently and to execute a task even in environments with disturbing moments. It is the cognitive domain perhaps most closely associated with everyday function (Royall, Palmer, Chiodo, & Polk, 2005) and, together with attention, often affected in neurocognitive as well as non-neurocognitive diseases. Difficulties with attention or executive function may for instance simulate memory problems due to decreased ability to stay focused on the task in individuals suffering from depression, one of the treatable differential diagnoses of dementia (Wright & Persad, 2007). Both capacities are also known to be affected by age (Harada et al., 2013) and a fair evaluation is dependent on reliable tools.

For a test to be reliable and valid in clinical practice without the help of a trained neuropsychologist, it should be easy to administrate and not too time consuming, and the effects of demographic factors such as gender, age and education should be known. The Trail Making Test A (TMT-A) (Reitan, 1955) measures attention and speed, and requires visual scanning, graphomotor speed and visuomotor processing speed as main cognitive functions. Gender is considered to have no or very little impact on the results, in contrast to education and age where high age and low education affect the results in a negative way (Llinas-Regla et al., 2017; Wei et al., 2018). The Symbol Digit Modalities Test (SDMT) (Smith, 1991) evaluates complex attention, visual scanning and tracking, motor speed and memory (Strauss, Sherman, & Spreen, 2006) and has been found to predict functional decline in aging (Classon, Fallman, Wressle, & Marcusson, 2016). SDMT is a fairly robust test, and even though some studies have found differences between men and women (Jorm, Anstey, Christensen, & Rodgers, 2004; Kiely, Butterworth, Watson, & Wooden, 2014), it has often been considered as not clinically relevant and different norms are not needed due to gender (Arango-Lasprilla et al., 2015; Peña-Casanova et al., 2009; Strauss et al., 2006). Many studies have however shown decreased test results with aging and higher points in well-educated subjects (Peña-Casanova et al., 2009; Richardson & Marottoli, 1996). The Victoria Stroop Test (VST) (Regard, 1981) is an executive test of selective attention and cognitive flexibility. The results from normative studies concerning impact of demographic variables such as gender and education are not consistent, but age is considered to have a negative impact on the results (Bayard, Erkes, & Moroni, 2011; Tremblay et al., 2016; Troyer, Leach, & Strauss, 2006). The Parallel Serial Mental Operations (PaSMO) (Nordlund et al., 2005) is an oral test of executive function and cognitive flexibility that does not require visual
scanning or psychomotor speed (Gothlin, Eckerstrom, Rolstad, Wallin, & Nordlund, 2017). To our knowledge, there are no published normative studies for PaSMO.

In the present study, community-dwelling elderly men and women were repeatedly tested at ages 85, 90 and 93 within the Elderly in Linköping Screening Assessment (ELSA 85). The aim of the study is to present normative values for the very old population for the four mentioned well-established tests and to investigate whether there are any age effects in the test results for the subgroup that participated throughout the study.

**Material and method**

*Participants and procedure*

Participants in this study were originally recruited for the longitudinal population study Elderly in Linköping Screening Assessment (ELSA 85). ELSA 85 initially invited all 85-year-olds (born in 1922) living in the municipality of Linköping in 2007 (n = 650 persons) to participate in the study. The study design is described in detail in Nagga, Dong, Marcusson, Skoglund, and Wressle (2012) and in Johansson, Marcusson, and Wressle (2019). 338 persons agreed to come to the Geriatric Clinic at Linköping University Hospital for an evaluation of cognitive function with neuropsychological assessments (described below) and for somatic check-up by a physician. Of the evaluated persons, 336 completed the Mini Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) and were included in this study.

Follow-ups were carried out with home visits when the participants were 90 and 93 years old. Only participants who had undergone all the previous stages of the study and completed the MMSE at each stage were included, n = 107 at 90 years of age and n = 51 at 93 years of age (Figure 1).

The assessments were administered by an occupational therapist, a nurse or a physician from the Geriatric Department at Linköping University Hospital and all the evaluated tests were part of a screening test battery, the Cognitive Assessment Battery (CAB) (Nordlund, Pahlsson, Holmberg, Lind, & Wallin, 2011). During all phases of the study, the participants could cancel their participation or decline to participate in one or more assessments and then continue with another test. Thus, not all participants have carried out all the tests. They also had the possibility to end the examination and continue at another time.

Information about the participants’ health was obtained from the somatic check-ups, the subjects themselves and their medical records.

*Exclusion criteria*

The focus of this study was to find normative values for the cognitively healthy population aged 85 and above. Therefore, persons with known diseases that can cause cognitive decline were excluded from the analysis. Exclusion criteria were as follows: severe head injury, traumatic cerebral bleeding, stroke, transient ischemic attack during the last six months, epilepsy, Parkinson’s disease, Parkinsonian disorders, Huntington’s disease, multiple sclerosis, normal pressure hydrocephalus, dementia, objective mild cognitive impairment, chronic obstructive pulmonary disease with hypoxia (saturation
less than 92% in air), insulin-dependent diabetes mellitus, generalized cancer, psychotic disorders, ongoing depression and drug abuse (Lezak, Howieson, Bigler, & Tranel, 2012). When detailed somatic information was not available, the participant was excluded.

To further minimize the risk of including persons with cognitive decline, due to somatic health problems not registered in their medical journal, persons who scored fewer than 24 points in the MMSE were excluded (Creavin et al., 2016). A recent study from Kvitting, Fallman, Wressle, and Marcusson (2019) suggests a cut-off of 26 points for persons aged 85 and above. However, changing the cut-off did not give any clinically relevant differences in results (results not shown).

**Cognitive tests**

The participants’ cognitive functions were evaluated with the Cognitive Assessment Battery (CAB) (Nordlund et al., 2011) and all the investigations were carried out in accordance with its
manual. The CAB consists of multiple assessments, some of them in established short versions. It was developed to investigate the cognitive domains of speed and attention, learning and episodic memory, visuospatial abilities, language, and executive functions. The tests evaluated in this study were executed in the following order: SDMT, VST, TMT-A and PaSMO.

The Trail Making Test (TMT) (Reitan, 1955) is one of the most widely used neuropsychological tests and consists of two subtests, TMT-A and TMT-B. In the CAB only TMT-A, the first subtest, is included. The participant is given the task of drawing a line between the numbers 1–25, scattered across a sheet, in the correct ascending order as quickly as possible. The score is the time in seconds to complete the task, with higher scores representing worse attentional capacity. No practice effects have been seen with long intervals (12 months) (Strauss et al., 2006).

The Symbol Digit Modalities Test (SDMT) (Smith, 1991) is a short, easily administered cognitive test which exists in both written and oral forms, of which the written form is used in the CAB. The participant is presented with a coding key of nine geometrical symbols, each representing a digit. After an initial 10 trial practice period, the subject is asked to recode as many symbols into digits as possible in 90 seconds using the key. The score is one point per correct digit, and the maximum result is 110 points. No significant practice effects have been noted over a two-year interval (Uchiyama et al., 1994).

The Victoria Stroop Test (VST) (Regard, 1981) is a shorter version of the Stroop test (Stroop, 1935). It consist of three subtests, each presenting a stimulus card with six rows of four items. The first test (VST1) shows 24 dots in the colors green, blue, red and yellow in a pseudorandom order. The second test (VST2) shows 24 common words, written in the four different colors. The third test (VST3) shows 24 color words (green, yellow, blue and red) written in one of the four colors, but so that the print color never corresponds to the color name (e.g. “blue” is written in green ink). For every subtest, the participant is asked to name the color of the dot, word or color word as quickly as possible and the examiner corrects any unnoticed errors. The score is the time in seconds, with higher scores indicating worse performance. There were no significant retest effects when middle-aged to older persons carried out a similar Stroop test with three-year intervals (Adolfsdottir, Wollschlaeger, Wehling, & Lundervold, 2017).

The Parallel Serial Mental Operations (PaSMO) (Nordlund et al., 2005) is a verbal test that is similar to the oral version of the TMT-B (Ricker, Axelrod, & Houtler, 1996) and the Alphanumeric Sequencing Test (Grigsby & Kaye, 1995), but includes all the letters in the alphabet with a reverse order between number and letter. It was originally developed at the Institute of Neuroscience and Physiology at the University of Gothenburg, Sweden. The participant is first given the training task of reciting the alphabet as quickly as possible (the Swedish version with 28 letters). Then the participant is asked to recite the alphabet together with the number that corresponds to the letter (e.g. A-1, B-2). The score for the test is the time in seconds, and if the participant makes a mistake the examiner helps by saying the last correct letter and number pair. Retest effects and practice effects on PaSMO or similar oral executive tests such as Oral TMT-B have not yet been established (Strauss et al., 2006).

**Ethical approval**

The ELSA 85 study, including permission to obtain data from all registers held by Östergötland County Council, has been approved by the Research Ethics Committee of
Linköping University, Sweden (2006: 141–06, 2012: 332–31; 2014: 455–31) and complies with the ethical rules for human experimentation stated by the Declaration of Helsinki (World Medical Association declaration of Helsinki, 1997). Written informed consent was collected from all participants at each step of the study.

**Statistical analysis**

Statistical analysis was carried out using SPSS version 24. Mean values and standard deviations were calculated for each age group and for each age divided by educational level: less than or equal to 9 years of schooling, or 10 years and above (initial analyses including three educational levels, 4–9 years, 10–12 years and 13 years and above, showed no difference between the two levels with higher years of schooling). Any differences between men and women and level of education were analyzed with independent sample t-tests. Age effects were analyzed with repeated measures for those individuals who had participated in all three test occasions (referred to below as Completers) and post hoc analyses between the age groups were carried out with the Bonferroni method. Sphericity was calculated using Mauchly’s test, assessing the hypothesis that the variance of the differences between conditions are equal, and when sphericity could not be assumed the significance was determined by correction of Greenhouse-Geisser (Field, 2018). The level of significance was set at 0.05.

Mean values and standard deviations for the whole ELSA 85 population, as well as demographic information about the participants excluded from this study, are presented in the supplementary material. Analyses or test results regarding the whole ELSA 85 population and the excluded participants will not be discussed further in this article.

**Results**

All the participants in this study, with known educational level, had at least four years of schooling. Descriptive information about the participants is shown in Table 1.

When comparing the 172 persons who dropped out or were excluded at either of the follow-ups for the 35 participants that were included at 93 years of age, it was noted that there were lower proportions of men (42% vs 60%) and participants living with family member (43% vs 63%) at baseline. The proportions with a high educational level

<table>
<thead>
<tr>
<th>Table 1. Descriptive information about the participants.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>114 (55)</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>93 (45)</td>
</tr>
<tr>
<td><strong>Living situation</strong></td>
</tr>
<tr>
<td>Living alone</td>
</tr>
<tr>
<td>111 (54)</td>
</tr>
<tr>
<td>Living with family member</td>
</tr>
<tr>
<td>96 (46)</td>
</tr>
<tr>
<td><strong>Residence</strong></td>
</tr>
<tr>
<td>Ordinary accommodation</td>
</tr>
<tr>
<td>203 (98)</td>
</tr>
<tr>
<td>Sheltered housing</td>
</tr>
<tr>
<td>4 (2)</td>
</tr>
<tr>
<td>Residential home</td>
</tr>
<tr>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Education (years)</strong></td>
</tr>
<tr>
<td>≤ 9</td>
</tr>
<tr>
<td>144 (70)</td>
</tr>
<tr>
<td>≥ 10</td>
</tr>
<tr>
<td>60 (29)</td>
</tr>
<tr>
<td>Missing data</td>
</tr>
<tr>
<td>3 (1)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
<tr>
<td>207</td>
</tr>
</tbody>
</table>

K. FÄLLMAN ET AL.
(28% vs 31%) and ordinary accommodation (98% vs 100%) were however almost the same in the two groups.

There was no difference between men and women for any test at any age group in this study except for the VST1 at age 85, where women had a lower mean (i.e. were faster) than men ($M_{\text{women}} = 18, SD = 5.7, M_{\text{men}} = 21, SD = 6.1, t(203) = -2.68, p = .008$).

The normative data for TMT-A, SDMT, VST and PaSMO are presented in Table 2. The norms are divided into 85-, 90- and 93-year-olds, as a total and in two educational levels. As can be seen, there was a significant difference due to educational level for all tests, except VST1, in the 85-year-olds. However, only performance on SDMT differed due to educational level at age 90, and at age 93 educational level had no significant impact on test results. It is worth noting that the drop-out frequency for PaSMO was higher than for the other three tests – 16% at 85 years, 32% at 90 years and 28% at 93 years – and that the drop-outs were mostly men or participants with low educational level.

As can be seen in Figure 2, TMT-A, SDMT, VST1 and PaSMO were all sensitive to aging with successively lower performance. However, there was no difference due to age for VST2 and VST3. The actual change over time as a mean was an increase of 11 seconds ($\pm 21$) for TMT-A, a decrease of 8 symbols ($\pm 6$) for SDMT, and increases of 3 seconds ($\pm 3$) for VST 1, 1 second ($\pm 7$) for VST 2, 4 seconds ($\pm 9$) for VST 3 and 19 seconds ($\pm 27$) for PaSMO. The Completers had an overall somewhat better result than the whole cognitively intact population.

**Discussion**

We present normative values for four easily administered cognitive tests that evaluate attention and executive function, derived from a large population of cognitively intact very old persons (i.e. over 85 years of age). The norms are presented as mean values and standard deviations to facilitate applicability in clinical practice. To the authors’ knowledge, this is the first study to present normative values for very old persons for PaSMO and the study with the highest number of nonagenarians for the other three tests. Even though there are other studies, e.g. Royall, Palmer, Chiodo, and Polk (2004), investigating attention and executive function in relation to high age, the present study is one of the largest normative studies measuring attention and executive function in persons 85 years of age and above.

As in various other normative studies, we found no significant gender differences in TMT-A or SDMT (Ashendorf et al., 2008; Peña-Casanova et al., 2009; Sheridan et al., 2006), or in PaSMO. In VST1 only, women were faster than men at the age of 85 years, but the difference was only 3 seconds and the clinical relevance was questionable. Tremblay et al. (2016) found that women outperformed men in all parts of VST, but their norms are based on younger elderly. Most other studies on VST and similar variants of the Stroop test do not recommend gender-specific norms (Strauss et al., 2006) that, in line with the present results, are not required for the oldest old.

We present our results divided into two educational levels, ≤9 and ≥10 years. For the octogenarians the results show an impact of educational level on performance, which is in accordance with previous studies (Ashendorf et al., 2008; Kiely et al., 2014; Tremblay et al., 2016). Surprisingly, there were no significant differences between high and low levels of education in the nonagenarians, except for SDMT at 90 years of age. This
Table 2. Normative data for the oldest old: Trail Making Test-A (TMT-A), Symbol Digit Modalities Test (SDMT), Victoria Stroop Test (VST) and Parallel Serial Mental Operations (PaSMO)

<table>
<thead>
<tr>
<th>Age</th>
<th>Edu</th>
<th>TMT-A</th>
<th>SDMT</th>
<th>VST 1</th>
<th>VST 2</th>
<th>VST 3</th>
<th>PaSMO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>M*</td>
<td>SD</td>
<td>p</td>
<td>n</td>
<td>M*</td>
</tr>
<tr>
<td>85</td>
<td></td>
<td>198</td>
<td>68</td>
<td>31</td>
<td></td>
<td>197</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>197</td>
<td>24</td>
<td>8</td>
<td></td>
<td>197</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>205</td>
<td>19</td>
<td>6</td>
<td></td>
<td>203</td>
<td>31</td>
</tr>
<tr>
<td>90</td>
<td></td>
<td>33</td>
<td>77</td>
<td>33</td>
<td></td>
<td>37</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>33</td>
<td>77</td>
<td>33</td>
<td></td>
<td>37</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>33</td>
<td>77</td>
<td>33</td>
<td></td>
<td>37</td>
<td>29</td>
</tr>
<tr>
<td>93</td>
<td></td>
<td>23</td>
<td>71</td>
<td>27</td>
<td></td>
<td>23</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>71</td>
<td>27</td>
<td></td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>71</td>
<td>27</td>
<td></td>
<td>22</td>
<td>20</td>
</tr>
</tbody>
</table>

* = seconds
** = numbers correct
Edu = years of education
*** = all participants, including three 85-year-olds and one 90-year-old with unknown educational level
contradicts most other studies, which have found educational level to be important for attention and executive function as measured by these tests. The lack of statistically significant differences could however be due to the sizes of the age groups in this study, with smaller groups at the very high ages. Mean performances show differences that are clinically significant (e.g. more than 5 seconds or 0.5 SD) for all ages in TMT-A, VST3 and PaSMO between the two educational levels. Bezdicek et al. (2017) found that age seemed to have a more pronounced effect compared to education in old age, but there are few studies with participants over 90 years of age with which to compare our data and more research needs to be done to be able to draw any conclusions about the present findings.

When scrutinizing the field for normative test results for the very old and TMT-A, the results are conflicting. Ashendorf et al. (2008) and Bezdicek et al. (2017), two fairly large studies (n = 117, 79–98 years of age and n = 104, 85–98 years of age, respectively) with many similarities to our study in the choice of participants, have two diverse normative results where our study results end up in between, together with Zalonis et al. (2008).

Our results for the SDMT are in consensus with the results from a normative study from Australia, Kiely et al. (2014), and with Richardson and Marottoli (1996), a smaller study often
referred to in clinical practice. There are other studies with both higher and lower norm values than those presented here, but they are from small samples or younger participants (Nielsen, Lolk, & Kragh-Sørensen, 1995; Vogel, Stokholm, & Jørgensen, 2013).

Only a few normative studies have been carried out for VST. In contrast to Bayard et al. (2011) and Tremblay et al. (2016), we chose not to focus on the interference score, but to calculate the time score to facilitate use in clinical practice. Our results are consistent with the results of Troyer et al. (2006) concerning VST3, but our participants have a higher score (i.e. take longer) for the VST1 (4–5 seconds longer mean) and VST2 (7–12 seconds longer mean), and our results are more diverse. The participants in Troyer et al. (2006) are however younger and have a higher educational level than the participants in this study.

To our knowledge, there are no published normative studies concerning PaSMO. The norms presented in the CAB manual are only applicable up to the age of 79, and our data suggest a higher and more diverse range for normative values for persons aged 85 years and above. With this in mind, and combined with the fact that there were many participants – especially those with low educational level and men – who did not complete this test, it may be questioned whether this is a clinically useful test for the very old population. Oral versions of alphanumerical tests such as the oral TMT-B and PaSMO are thus likely to be too challenging for the oldest old and not suitable alternatives to pen-and-pencil formats in elderly people with disabilities, disproportionately penalizing low-educated men. It is however important to note that PaSMO is the last test in the CAB, which may be a contributing factor to the drop-out frequency. More normative research and validation for PaSMO and similar oral tasks need to be done to investigate the clinical use of these executive tests.

When analyzing the results from the Completers, there were declines in results due to age in all the tests except for those that involve inhibition (VST2 and VST3). These results are in general accordance with previous studies showing age-related declines in cognitive speed, attention and executive functioning (Royall et al., 2004; Salthouse, 1996, 2016). The lack of aging effects in inhibition conflicts with some (Adolfsdottir et al., 2017; Troyer et al., 2006; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006) but not all (Verhaeghen & De Meersman, 1998) previous research. It has been suggested that inhibition per se is actually well preserved during aging and worsening Stroop performance, when seen, is more a reflection of cognitive slowing.

There are some limitations in the study design that needs to be highlighted. The definition of “cognitively intact” is an estimation done after the data were collected. This can lead both to the risk of including persons with preclinical cognitive disease, but also to excluding cognitively healthy participants with e.g. somatic disease, leading to a biased test result. The careful exclusion of participants with conditions known to impact on cognition helps to ensure that the norms presented reflect healthy aging, but as we did not have the information about e.g. biomarkers such as beta amyloid and tau proteins the definition is an estimation. Normative studies in the oldest old are scarce and for natural reasons the decrease in number of participants in the highest age groups contribute to possible population differences between ages, like survival of the fittest. This could be an explanation to the somewhat unexpected result of similar mean test results for the evaluated tests throughout the eight years of the study, where the healthier and more cognitively intact participants are more likely to continue to participate in the study, resulting in a more homogenous group throughout the years of the study. Therefore, the age effect was studied
in the subgroup that participated in all time points, reducing the possibility of differences in population between time points. The normative data for the nonagenarians in this study are based on small sample sizes. This is not ideal for normative data because of the risk of sampling errors and more normative research with larger samples needs to be done for the oldest old population to confirm our results. However, for a certain very old patient, the normative data based on a small demographically fit sample may by more representative than data based on a large age heterogeneous sample and therefore the norms presented in this study are nonetheless useful in clinical practice. Where comparisons can be made, the present normative results are in accordance with those of previous studies.

**Conclusion**

The normative values presented in this study are valid and representative for very old persons in Sweden and other countries with similar social contexts and educational systems. The norms are useful in the clinical evaluation of attention and executive function, such as in the investigation of neurodegenerative disease, and the decline in test results due to age highlights the need for updated normative values for this age group.

**Acknowledgments**

We would like to thank the participants for devoting their time and energy to this project. We would also like to thank our colleagues who have taken part in the data collection with enthusiasm and endurance.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This study was performed by the authors while employed by Linköping University and/or Östergötland County Council. Support staff (nurse, occupational therapist) were funded by a research grant from Linköping University (number LIO-696631, LIO 602761 and LIO 537591). The funders had no input into or influence on the study; Linköpings Universitet [LIO- 696631, LIO-602761, LIO- 537591].

**Contributors**

KF: drafting the manuscript, analysis and interpretation of the data, acquisition of data and statistical analysis. LL: analysis and interpretation of data, revising the manuscript. EW: recruitment of the study cohort, study concept and design, supervision of data collection and interpretation of data, and obtaining funding. JM: recruitment of the study cohort, study concept and design, supervision of data collection and interpretation of data, revising the manuscript and obtaining funding. EC: drafting/revising the manuscript, study concept and design, analysis and interpretation of the data, study supervision and obtaining funding. All authors have provided critical comments on the manuscript and approved the final version.
Data sharing statement

Data cannot be made publicly available. Authorized researchers who have ethics approval and an institutionally approved study plan can apply for data sharing through the Department of Clinical and Experimental Medicine, Linköping University. For more details, please contact the corresponding author.

Impact statement

We certify that this work is recent novel clinical research. It presents useful clinical data on the use of Trail Making Test A, the Symbol Digit Modalities Test, the Victoria Stroop Test and Parallel Serial Mental Operations, for persons aged 85 years and above.

ORCID

Katarina Fällman http://orcid.org/0000-0002-5289-8831

References


